

TARTU ÜLIKOOL
LOODUS- JA TÄPPISTEADUSTE VALDKOND
MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT
BIOTEHNOLOOGIA ÕPPETOOL

**Ravimivastuses oluliste tsütokroom P450 geenide varieeruvuse
määramine Eesti populatsioonis**

Bakalaureusetöö

12 EAP

Karin Kõrgesaar

Juhendajad MSc Kristi Krebs

PhD Lili Milani

TARTU 2018

Ravimivastuses oluliste tsütokroom P450 geenide varieeruvuse määramine Eesti populatsioonis

Bakalaureusetöö analüüsib ravimi lagundamises osalevate oluliste tsütokroom P450 geenide alleeli sagedusi ja nende põhjal ennustatavaid metabolismi fenotüüpe Tartu Ülikooli Eesti Geenivaramu täisgenoomi andmetest. Kuna *CYP2D6* geeni genotüpiseerimine on keerukas väga polümorfse lookuse tõttu, siis määrati antud töös ka geenile koopiaarvu variatsioonide sagedusi. Selle põhjal näitab käesolev töö, miks oleks Eestis vajalik patsientide geneetiline testimine enne ravimi välja kirjutamist arsti poolt.

Märksõnad: farmakogenoomika, star-alleel, tsütokroom P450, ADR, CNV, CPIC, PharmGKB

CERCS: B220

Identification of genetic variation in drug metabolizing cytochrome P450 genes in Estonian population

The Bachelor's thesis analyses allelic frequencies of cytochrome P450 genes, contributed in drug metabolism, and predicted functionality across University of Tartu Estonian Genome Centre gene donors. Because *CYP2D6* gene is highly polymorphic and therefore challenging to genotype, this research also detects frequencies of *CYP2D6* copy number variations for correct interpretation of alleles. Based on this data, this thesis exhibits clear confirmation of the necessity to make pre-treatment testing to optimize drug therapy in Estonia.

Keywords: pharmacogenomics, star-allele, cytochrome P450, ADR, CNV, PharmGKB

CERC: B220

SISUKORD

KASUTATUD LÜHENDID	4
SISSEJUHATUS.....	5
1. KIRJANDUSE ÜLEVAADE.....	6
1.1 Farmakogeneetikast farmakogenoomikani.....	6
1.2 Ravimi metabolism.....	7
1.2.1 Faas I reaktsioonid.....	8
1.2.2 Faas II reaktsioonid	9
1.2.3 <i>CYP2C9, CYP2C19, CYP3A5, CYP4F2</i>	10
1.2.4 <i>CYP2D6</i>	11
1.3 Farmakogeneetilised uuringud.....	14
1.4 Farmakogeneetilistest uuringutest kinnitatud geen-ravim seosteni	15
1.4.1 PharmGKB	15
1.4.2 CPIC	17
2. UURIMUS.....	19
2.1 Töö eesmärk	19
2.2 Materjalid ja metoodika	19
2.2.1 Star-alleelide määramine	19
2.2.2 <i>CYP2D6</i> koopiarvud.....	21
2.3 Tulemused	22
2.3.1 Star-alleelide määramine	22
2.3.2 <i>CYP2D6</i> koopiarvud.....	28
2.4 Arutelu.....	33
KOKKUVÕTE	39
SUMMARY	40
TÄNUAVALDUSED.....	42
KASUTATUD KIRJANDUS.....	43
KASUTATUD VEEBIAADRESSID	48
LISAD	49
LIHTLITSENTS.....	53

KASUTATUD LÜHENDID

ADMET	Ravimi imendumine, jaotumine, metabolism, ekskretsioon, toksilisus (<i>Absorption, Distribution, Metabolism, Excretion, Toxicity</i>)
ADR	Ravimi kõrvaltoime (<i>adverse drug reaction</i>)
CA	Kliinilised annotatsioonid (<i>clinical annotations</i>)
CNV	Geeni koopiarvu variatsioonid (<i>genetic copy number variations</i>)
CPIC	<i>Clinical Pharmacogenetics Implementation Consortium</i>
CYP	Tsütokroom P-450
EM/NM	Normaalne metaboliseerija (<i>extensive/normal metabolizer</i>)
G6PD	Glükoos-6-fosfaat dehüdrogenaas
GWAS	Ülegenoomne assotsiatsiooni-uuring (<i>genome-wide association studies</i>)
HGP	Inimese genoomi projekt (<i>Human Genome Project</i>)
IM	Keskmise metaboliseerija (<i>intermediate metabolizer</i>)
NGS	Teise põlvkonna sekveneerimine (<i>next generation sequencing</i>)
NTI	Kitsa terapeutilise aknaga ravimid (<i>narrow therapeutic index drugs</i>)
PharmGKB	<i>The Pharmacogenomics Knowledge Base</i>
PM	Aeglane metaboliseerija (<i>poor metabolizer</i>)
RM	Kiire metaboliseerija (<i>rapid metabolizer</i>)
SNP	Ühenukleotiidiline polümorfism (<i>single nucleotide polymorphism</i>)
SNV	Ühenukleotiidne variatsioon (<i>single nucleotide variation</i>)
TPMT	Tiopuriin-S-metüültransferaas
TÜ EGV	Tartu Ülikooli Eesti Geenivaramu
UM	Ülikiire metaboliseerija (<i>ultrarapid metabolizer</i>)
VA	Variatsioonide annotatsioonid (<i>variant annotations</i>)
WES	Kogu eksoomi sekveneerimine (<i>whole exome sequencing</i>)
WGS	Kogu genoomi sekveneerimine (<i>whole genome sequencing</i>)

SISSEJUHATUS

Üks ja sama ravim, mida arstid erinevatele patsientidele kirjutavad, ei toimi alati kõigile ühtemoodi. Mõnele ei mõju ravim efektiivselt, teisel juhul esineb olukordi, kus ravimil on toksiline mõju. Tulemuseks võivad olla ohtlikud kõrvalmõjud, mis halvimal juhul võivad lõppeda ka surmaga. Paljudel ravimitel (nt digoksiin, varfariin, takroliimus jt) on kitsas terapeutiline aken (*narrow therapeutic index drugs*, NTI), mis viitab sellele, et piir ravimi efektiivsuse ja toksilisuse vahel on küllaltki õhuke (Jiang *et al.*, 2015). Norra haiglates läbi viidud uuringust selgus, et 35% hospitaliseeritud patsientide raviks kasutati just NTI ravimeid (Blix *et al.*, 2010).

On mitmeid faktoreid, mis mõjutavad ravimitoime varieeruvust, nt keskkond, toitumine, teised kasutatavad ravimid ja elustiil. Olulisimaks mõjuteguriks peetakse siiski indiviidide geneetilist varieeruvust. Tänapäeval teostatavad genoomiuuringud on oma olemuselt väga põhjalikud, mis annab võimaluse tuvastada geneetilist varieeruvust korraga mitmetes ravimivastuses olulistes geenides ning avastada ka uusi variatsioone. Üks ravimivastuses oluline geeniperekond on *CYP* geenid. Geneetilised variatsioonid *CYP* geenides võivad mõjutada ravimi metabolismis oluliste ensüümide funktsionaalsust, mis tulemusena võib mõjutada ravimi efektiivsust või toksilisust. Informatsioon ravimivastuses oluliste geenide varieeruvuse kohta aitaks ka muuta meditsiini personaalsemaks ning ravimite väljakirjutamist inimese genoomist lähtuvaks. See võimaldaks tõsta ravimite efektiivsust, potentsiaalselt vältida kõrvaltoimeid ning leida ka igale indiviidile täpsemaid doosisoovitusi iga ravimi jaoks.

Seetõttu on antud bakalaureusetöö eesmärk uurida Tartu Ülikooli Eesti Geenivaramu andmete põhjal *CYP* perekonna geenide varieeruvuse sagedust Eesti populatsioonis ning võrrelda tulemusi Euroopa populatsioonidega ning lisaks kirjeldada ja ka valideerida olulisima *CYP* perekonnaliikme *CYP2D6* koopiaarvude esinemist.

1. KIRJANDUSE ÜLEVAADE

1.1 Farmakogeneetikast farmakogenoomikani

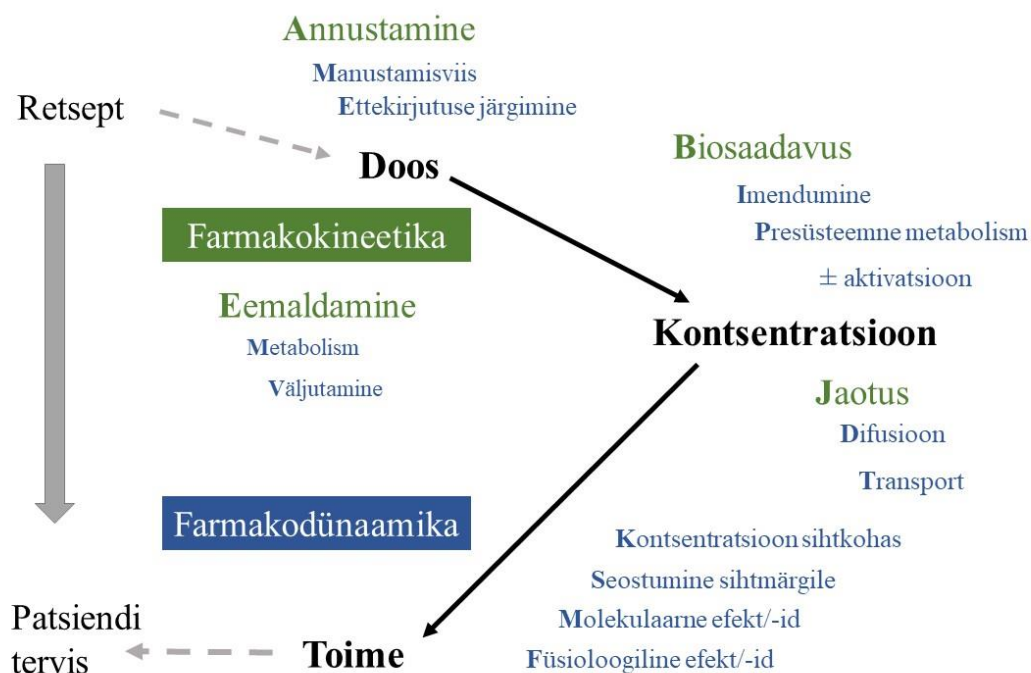
Inimesed erinevad üksteisest nii füsioloogiliste tunnuste (vanus, sugu, populatsiooniline päritolu), elustiili (toitumine, suitsetamine, kokkupuude toksiliste ainetega) kui ka geneetilise informatsiooni poolest. Kõik need faktorid on tihedalt seotud organismi võimega tulla toime erinevate haigustega ning põhjustavad varieeruvust ravimivastuses (Pirmohamed, 2014). Farmakogenoomika on teadusharu, mille eesmärgiks on selgitada, kuidas indiviidide geneetilised variatsioonid mõjutavad organismis ravimite efektiivsust ja toksilisust. Interindividaalset varieeruvust ravimi toimes põhjustavad geneetilised variatsioonid nagu ühenukleotiidsed polümorfismid (*single nucleotide polymorphism*, SNP), koos päranduvate geneetiliste variantide (nt SNP-de) kogumid ehk haplotüübid, deletsioonid, insertioonid ning geeni koopiarvu variatsioonid (*genetic copy number variations*, CNV). Geneetiline polümorfism ravimi imendumist, jaotumist, metabolismi, väljutamist ning toksilisust (*Absorption, Distribution, Metabolism, Excretion, Toxicity*, ADMET) reguleerivates geenides võib tulemusena põhjustada erinevatel patsientidel samale ravimile erisuguseid toimeid. Ravimiannus, mis ühele on efektiivne võib teisel patsiendil tekitada ohtlikke kõrvalmõjusid (*adverse drug reactions* ADR) või halvimal juhul lõppeda surmaga. Sellest tulenevalt on farmakogenoomikal oluline osa personaalmeditsiini arendamises, mille ideeks on välja töötada tõhus ja ohutu raviplaan igale patsiendile vastavalt tema geneetilistele iseärasustele (Ahmed *et al.*, 2016).

Arvatakse, et Pythagoras oli esimene, kes kirjeldas farmakogeneetikat. Ligikaudu aastal 510 eKr pani Pythagoras tähele, kuidas osadel inimestel tekkis hemolüütiline aneemia pärast põldubade söömist, teistel aga mitte. Põhjus avastati kaks sajandit hiljem - defekt glükoos-6-fosfaadi dehüdrogenaas (G6PD) geenis mõjutab hemolüütilise aneemia esinemist kokkupuutel põldubadega (Nebert, 1997; Alving *et al.*, 1956). Tõsisemalt hakati farmakogeneetikaga tegelema 1960. aastate alguses, kui viidi läbi mitmeid kliinilisi katseid, uurimaks, kuidas geenid mõjutavad ravimitoimet (Motulsky, 1957; Kalow, 1962). Saksa geeniteadlane Friedrich Vogel võttis esimesena kasutusele mõiste farmakogeneetika aastal 1959 (Vogel, 1959). Kõigest kuus aastat hiljem avastasid Watson ja Crick DNA struktuuri, mis aitas mõista, kuidas geenid kontrollivad keemilisi protsesse ning pani aluse DNA sekveneerimisele (Watson and Crick, 1953). Oluliseks verstapostiks farmakogeneetikas peetakse ka maksaensüümi geeni *CYP2D6* polümorfismi avastamist. *CYP2D6*

oli esimene inimese ravimimetabolismi geen, mis esmakordselt klooniti ja kirjeldati aastal 1987 (Gonzalez *et al.*, 1988). Mitmed hüpertensioonivastaste ravimite (nt debrisoikiin) ainevahetushäirete uuringud näitasid, et *CYP2D6* funktsionaalse toime puudumine mõjutab umbes 60 erineva ravimi toimet (Mahgoub *et al.*, 1977; Marez *et al.*, 1997; Kalow, 2006). Järgnev teaduse ja tehnoloogia areng võimaldas paremini mõista, kuidas genoomi keerukus mõjutab inimese tervist ja haiguse arengut. Kui algselt arvati, et ainult üksikud geenid avaldavad mõju ravimi imendumisele ja ravimivastusele, siis pärast inimese genoomi projekti (*Human Genome Project*, HGP) valmimist 2003. aastal veenduti, et haiguste teket ja ravimite metabolismi mõjutavad koos mitmed geneetilised ja keskkonna tegurid (Maggo *et al.*, 2016; Daly, 2010). Ülegenoomsed uuringud ja kiire areng sekveneerimismeetodite vallas on aidanud kaasa ka farmakogenoomika mõiste arenemisele. Kuigi esmakordselt mainiti seda kirjanduses juba 1997. aastal, siis algselt kasutati farmakogeneetikat ja –genoomikat sünonüümidenä. Alles hiljem, kui mõisteti, et mitme geeni koostoime mõjutab kindla metabolismi fenotüübi arengut, siis hakati rohkem kasutama farmakogenoomika terminit (Meyer, 2004; Daly, 2010).

1.2 Ravimi metabolism

Umbes 10% turul olevatest ravimitest on algselt eelravimi (*prodrug*) kujul, millel puudub farmakoloogiline toime sihtmärgile. Eelravimite kasutamise eesmärgiks on tõsta farmakoloogiliselt aktiivsete ainete lahustuvust, imendumist, keemilist stabiilsust ning ennetada nende varast metabolismi ning toksilisust organismis. Selleks et saavutada ühendi soovitud metaboolne aktiivsus, tuleb need ühendid muundada aktiivseteks metaboliitideks (Zawilska, *et al.*, 2013). Lipofiilsed ravimid tuleb muuta vees lahustuvaks, et neid oleks võimalik organismist väljutada. Vastasel korral võib ravim organismis kuhjuda ning muutuda toksiliseks. Ravimi metabolismi läbi viivad keemilised reaktsioonid toimuvad kahes faasis - faas I ja II. Ensüümid, mis osalevad ravimite metabolismis, leidub peamiselt seedetraktis (maks, peensool, jämesool), aga ka neerudes, ajus, rinnakus ning kopsudes. Tavaliselt paiknevad ensüümid raku sisemembraanil ja tsütosoolis. Neid ensüüme leidub nii prokarüootides, eukarüootides, taimedes, seentes, bakterites, arhedes kui ka viirustes (Gopisankar, 2017). Geneetilised variatsioonid neid ensüüme kodeerivates geenides võivad mõjutada suuresti ravimite farmakokineetikat ja farmakodünaamikat. Farmakokineetika uurib, kuidas ravimit lagundatakse ning farmakodünaamika keskendub ravimi toimemehhanismidele (retseptorid,ioonkanalid, ensüümid) (Joonis 1) (Pirmohamed, 2014).



Joonis 1. Retsepti väljakirjutamisest kuni patsiendi terviseni. Farmakokineetika ja -dünaamika on peamised tegurid, mis mõjutavad individuaalset ravimivastust. Illustratiivne joonis näitab, millised on olulisemad etapid, mis mõjutavad ravimivastust organismis (Doogue ja Polasek, 2013). (Tõlgitud ja kohandatud).

1.2.1 Faas I reaktsioonid

Esimeses faasis muundatakse ravim vees lahustuvamaks või aktiivsemaks ühendiks. Toimuvad oksüdatiivsed, reduktiivsed ning hüdrolyütilised reaktsioonid. Faas I metabolismi eest vastutavad peamiselt tsütokroom P450 (CYP) valkude superperekonna oksüdatiivsed ensüümid, mis paiknevad hepatotsüütide siledapinnalises endoplasmaatilises retiikulumis (ER) (Ozdemir *et al.*, 2002). Inimese genoomi projektis identifitseeriti 57 geeni, mis kodeerivad CYP ensüüme. Umbkaudu 12 erinevat CYP perekondade 1, 2 ja 3 hulka kuuluvatest ensüümidest reguleerivad ravimimetabolismi. Variatsioonid CYP perekonna valke kodeerivates geenides põhjustavad indiviididel erinevaid metabolismitüüpe ning mõjutavad sellega ravimivastust (Zanger ja Schwab, 2013). Tsütokroome kodeerivad geenid on väga polümorfseid. Polümorfismid tekitavad geenis duplikatsioone, deletsioone, aminohappelisi muutusi ning enneaegseid stopp-koodoneid. Mitmetes teadustöodes on näidatud, et 57 CYP geenis on leitud üle 6000 ühenukleotiidse variatsiooni (*single nucleotide variation*, SNV). Nendest 1445 varianti on leitud 12 inimese CYP geenis, mis

reguleerivad ravimimetabolismi (Fujikura *et.al.*, 2015). CYP variatsioonide defineerimiseks kasutatakse farmakogeneetikale omast star-alleel (*star-allele* (*)) nomenklatuuri, kus määratakse kindlad haplotüübid, mis kirjeldavad ühe või enama variatsiooni esinemist. *1 määrab tavaliselt referents haplotüübi ehk tavapärase funktsiooniga alleeli. *1 on harilikult esimene leitud järjestus, mis kodeerib funktsionaalset valguprodukti, kuid mis alati ei ole kõige levinum star-alleel erinevates etnilistes populatsioonides. Referents haplotüübiga võrreldakse iga järgnevat leitud järjestust ning tuvastatakse teisi polümorfseid haplotüüpe (Kalman *et al.* , 2016). Erinevate star-alleelide esinemise järgi CYP-e kodeerivas geenis klassifitseeritakse individidid seejärel fenotüübi järgi aeglasteks (*poor metabolizers*, PM), keskmisteks (*intermediate metabolizers*, IM), tavapärasteks (*extensive metabolizers*, EM or *normal metabolizer*, NM), ülikiireteks (*ultrarapid metabolizer*, UM) ning mõnel juhul ka kiireteks (*rapid metabolizer*, RM) metaboliseerijateks. PM viitab kahele mittefunktsionaalsele alleelile, mis tähendab, et tulemusena kodeeritakse inaktiivne ensüüm. IM korral on tegu ühe funktsionaalse alleeliga ja ühe defektse või kahte osaliselt defektse alleeliga, mille tulemuseks on tavapärasest veidi aeglasem metabolism. EM puhul on tegemist tavapärase fenotüübiga ning see on ka kõige sagedamini esinev fenotüüp erinevates populatsioonides. UM korral esineb kas geeni duplikatsioon või valgu funktsionaalsust tõstev alleel, mille tulemusena ravimi metabolism toimub tavapärasest kiiremini (Gopisankar, 2017; Ahmed *et al.*, 2016). Tavapärasest erineva fenotüübi korral võivad individidid seetõttu vajada kohandatud annusesoovitusi. Näiteks on teada, et patsiendid, kes kannavad kahte inaktiivset alleeli *CYP2D6* geenis (*3-*8, *11-*16, *19-*21, *38, *40, *42) ja on seetõttu aeglased metaboliseerijad, PM, peaksid saama tavapärasest annusest 50% väiksema antidepressandi klomipramiini annuse (Swen *et al.*, 2011). Kuna erinevates populatsioonides esineb erinevusi ravimivastuses oluliste alleelide sageduses, siis oleks oluline uurida, kui sagedased on tavapärasest erineva metabolismiga fenotüübid konkreetses populatsioonis, et teada, kes potentsiaalselt vajaksid ka kohandatud annuseid.

1.2.2 Faas II reaktsioonid

Faas II metabolism hõlmab tekkinud metaboliitide detoksifikatsiooni ning inaktiveerimist. Teise faasi reaktsioonensüümideks on enamasti konjugatiivsed transferaasid (nt UDP-glükuronosüültransferaas (UGT), sulfotransferaas (SULT), N-atsetüültransferaas (NAT), glutatioon S-transferaas (GST), tiopuriin-S-metüültransferaas (TPMT) jt. Nende ensüümide abil toimub metabolismiproduktide väljutamine organismist uriini ja sapi vahendusel. Mõnel juhul võivad

geneetilised mutatsioonid faas II ensüüme kodeerivates geenides või teised samaaegselt tarvitavad ravimid suruda maha ensüümi funktsionaalsuse. Sel põhjusel ei muudeta metaboliite enam inaktiivseteks, vaheühendid hakkavad kuhjuma ning tekitavad toksilisi reaktsioone, mis põhjustavad ADR-e (nt maksakahjustusi) (Rowland *et al.*, 2013; Yiannakopoulou, 2013).

1.2.3 *CYP2C9*, *CYP2C19*, *CYP3A5*, *CYP4F2*

CYP2C9 kodeeriv geen asub kromosoomi 10q24.2 lookuses. Inimese maksa mikrosoomi ensüümidest moodustab *CYP2C9* 18%. Selle ensüümi poolt metaboliseeritavate ravimite hulka kuuluvad põletikuvastased ravimid, diabeediravimid, antikoagulandid ning epilepsiaravimid. Kõige olulisemateks *CYP2C9* alleelideks peetakse star-alleele *2 ja *3, mis vähendavad *CYP2C9* aktiivsust (Ahmed *et al.*, 2016).

Polümorfne *CYP2C19* geen asub kromosoomis 10q24. Selle geeni poolt kodeeritud ensüümide vahendusel metaboliseeritakse peamiselt ärevushäirete ravimeid, prootonpumba inhibiitoreid, epilepsiaravimeid ning malaariavastaseid biguaniide (Zhou *et al.*, 2008). Ravimite metabolismi mõjutavad enim *CYP2C19* star-alleelid *2 ja *3, mille kandjad on tavapärasest aeglasemad metaboliseerijad (Ahmed *et al.*, 2016).

Geen *CYP3A5* on üks kolmest *CYP3A4* isovormist, mis kõik paiknevad kromosoomis 7q21-q22.1. Kuivõrd *CYP3A5* ja *CYP3A4* aminohappeliste järjestuste identsus on suurem kui 85%, siis metaboliseerivad need ensüümid ka samu ravimite substraate. Tüüpilised ravimid, mida metaboliseeritakse on immunosupressandid (Zanger ja Schwab, 2013). Funktsionaalseid *CYP3A5* geeni alleele ekspresseeritakse peamiselt Aafrika populatsioonides. Europiidses populatsioonis on sagedaimaks star-alleeliks mittefunktsionaalne alleel *3 (Ingelman-Sundberg *et al.*, 2007).

Üks seni kõige vähem uuritud *CYP* geene on kromosoomi 19q13 lookuses paiknev *CYP4F2*. *CYP4F2* on peamine vitamiin K oksüdaas maksas. Siiani on seda ensüümi peamiselt seostatud ühe olulise vere hüübimist vähendava ravimi, varfariini, metabolismiga. Star-alleel *3, mis mõjutab varfariini metabolismi, on eelkõige levinud han hiinlaste seas ning ka mõnes Euroopa populatsioonis (www.pharmgkb.org/vip/PA166169424, külastatud 26. mai 2018).

1.2.4 CYP2D6

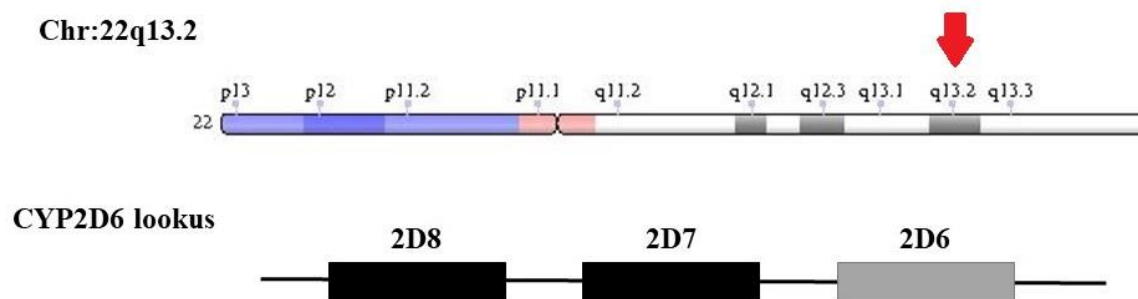
CYP2D6 kuulub tsütokroom P450 alamperekonda 2, sub-perekonda D ja on liige 6. See on kõige rohkem uuritud *CYP* geen, mis teadaolevalt mõjutab 25% turul olevate ravimite metabolismi (Zhou, 2009). *CYP2D6* kodeerib ensüümi, mis on oluline erinevatesse terapeutilistesse ravimiklassidesse kuuluvate ravimite metabolismis ja bioaktivatsioonis (nt antidepressandid, antipsühhootikumid, valuvaigistid, kõhavastased vahendid, beetablokaatorid, antiarütmikumid jt) (Zhou, 2009; Zanger ja Schwab, 2013; Gaedigk, 2013). *CYP2D6* geen on väga polümorfne. Sellel on leitud üle 100 erineva star-alleeli variandi (Tabel 1). Enamus geeni star-alleelidest on määratud kas ühe või enama SNP-i või insertsoonide-deletsioonide poolt, aga ka CNV kaudu, mis on *CYP2D6* geeni deletsiooni või duplikatsiooni tagajärg.

Tabel 1. *CYP2D6* alleelide funktsionaalsuse klassifikatsioon (Ahmed *et al.*, 2016; Gopisankar, 2017).

Mittefunktsionaalsed alleelid	*3 *4 *5 *6 *7 *8 *10 *11 *12 *13 *14 *15 *16 *18 *19
(PM)	*20 *21 *38 *40 *42 *44 *56 *62 *68 *92 *101
Osalise funktsiooniga alleelid	*10 *14 *17 *18 *36 *41 *47 *49 *50 *51 *52 *54 *55
(IM)	*57 *59 *69 *72
Suurenenud aktiivsusega	
alleelid (UM)	*2A *17x2 *27 *35 *39 *41x2 *48

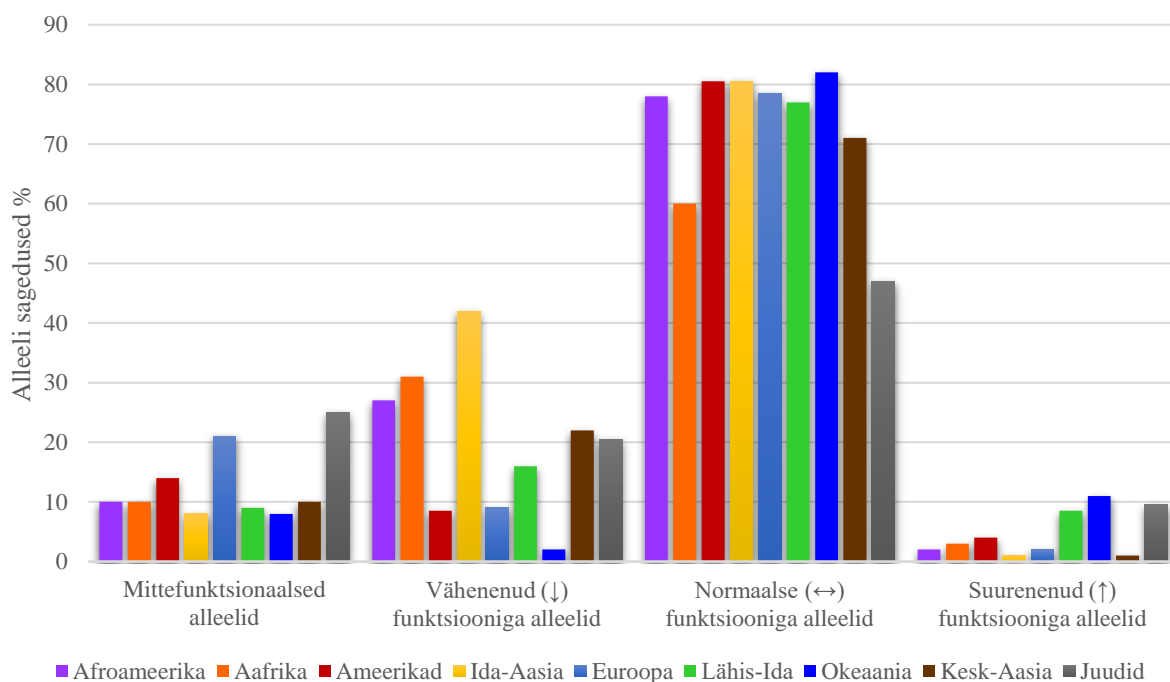
CYP2D6 on kromosoomi 22q13.1 lookuses asuv 45 kilobaasi suurune geen. Geenilookus on väga kompleksne ning sisaldab lisaks *CYP2D6* geenile veel kaht eraldiseisvat pseudogeeni *CYP2D7* ja *CYP2D8* (Joonis 2). *CYP2D6* on neist ainuke valke kodeeriv geen (Kimura *et al.*, 1989). Kõik kolm geeni sisaldavad 9 eksonit. *CYP2D7* ja *CYP2D6* on identsete või peaaegu identsete regioonidega ning omavad mitmeid homoloogilisi järjestusi. See on üks põhjuseid, miks *CYP2D6* geenuuringud on raskendatud. Need kaks geeni võivad moodustada hübriidgeene, mis tulemusena sisaldavad väikeseid segmente mõlemast geenist. Hübriidgeenid tekivad suurte deletsioonide tagajärjel, kus geeni *CYP2D7* 5' segment ühineb *CYP2D6* 3' osaga. Sellised variatsioonid ei pruugi alati ensüümi

funktsiooni mõjutada, kuid on leitud, et mõnel juhul võib ensüümi funktsionaalsus nii suurenda kui ka täielikult kaduda. Kui *CYP2D7/2D6* hübriidgeenis on *CYP2D7* ekson 1, siis loetakse need geenid mittefunktsionaalseteks, sest eksonis on T-insertsioon, mis põhjustab lugemisraami nihkeid ning varast translatsiooni terminatsiooni (Sim *et al.*, 2012; Gaedigk, 2013). Lisaks raskendab pseudogeenide *CYP2D8* ja *CYP2D7* ning *CYP2D6-CYP2D7* hübriidgeenide, näiteks *CYP2D6*36*, olemasolu *CYP2D6* geeni koopiarvude määramist. CNV-de põhjustatud ümberkorraldused tekitavad omakorda raskusi *CYP2D6* genotüpiseerimisel, sekveneerimisel ja fenotüübi ennustamisel. Põhjus on selles, et kõik CNV-d ei põhjusta samasugust ensüümi funktsionaalsust, s.t alleeli duplikatsioonid ei viita alati suurenenud ensüümi funktsioonile. Näiteks *CYP2D6*17xN* või **36xN* esinemine põhjustab ensüümi funktsiooni vähenemist ning *CYP2D6*4xN* korral puudub ensüümil funktsioon täielikult, kuna indiviidil esineb sel juhul mittefunktsionaalne *CYP2D6* koopia. Vähenenud aktiivsust põhjustab terve geeni deletsioon (*CYP2D6*5*) ning suurenenud aktiivsust >2 funktsionaalse koopia olemasolu (**1xN*, **2xN*) (Ramamoorthy *et al.*, 2010; Yang *et al.*, 2017). Vaatamata pidevalt arenevatele testimisplatvormidele, leidub *CYP2D6* lookuses ikka veel seni avastamata haruldasi ja populatsiooni spetsiifilisi variatsioone, mis piiravad täpset *CYP2D6* geeni analüüsi, mistõttu tuleks suhtuda tulemuste tõlgendamisse kriitiliselt (Gaedigk, 2013; Yang *et al.*, 2017).

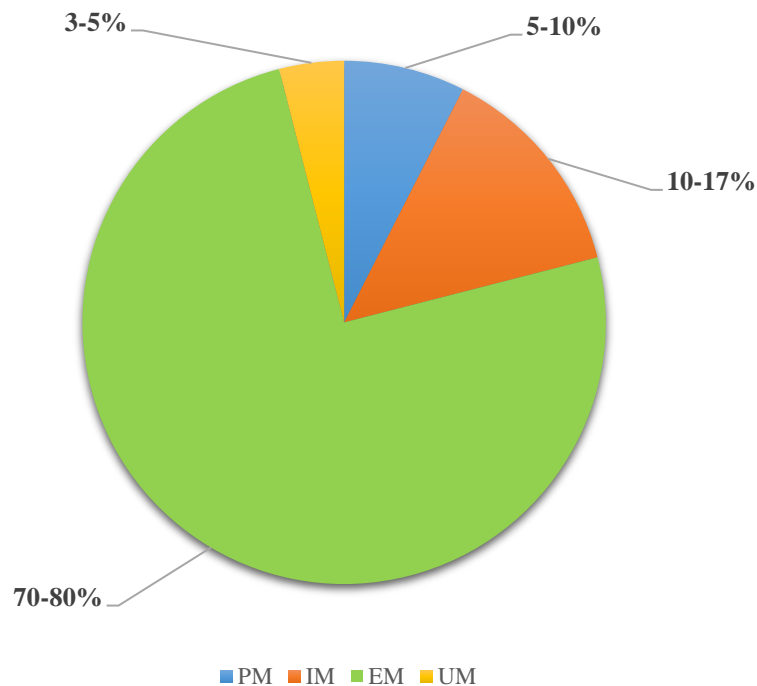


Joonis 2. *CYP2D6* asukoht kromosoomis ja geenilookus. (Kromosoomi joonis modifitseeritud www.ghr.nlm.nih.gov/gene/CYP2D6#location põhjal, külastatud 24. mai 2018)

CYP2D6 geeni alleelide põhjal ennustatud fenotüüpide esinemissagedus varieerub populatsioonide vahel. On leitud allelele, mis esinevad erinevates populatsioonides sama sagedusega, kuid on ka variante, mille sagedused erinevad oluliselt või mida leidub ainult kindlates etnilistes gruppides. Alleelisageduste põhjal selgub, et PM fenotüüpi leidub enam juudi (6%) ning europiidses populatsioonis (Euroopa ja Põhja-Ameerika) (keskmiselt 5,4%) ning kõige vähem Ida-Aasias, Kesk-Aasias, Okeaania ja Lähis-Idas (alla 1%). UM fenotüüpi esineb kõige enam Okeaania (21,2%), aškenazi juutidel (11,5%) ning Lähis-Idas (11,2%). Märkimisväärselt suurem varieeruvus populatsioonide vahel esineb ka IM fenotüübi sagedustes, vaheldudes 2,8%-st (Kesk- ja Lõuna-Ameerikas) kuni 11%-ni (aafriklaste ja afroameerika populatsioonid) (Joonis 3) (Gaedigk *et al.*, 2017). Europiidses populatsioonis on kõige enam levinud fenotüüp EM (70-80%) ning kõige vähem on UM (3-5%) (Joonis 4) (Sachse *et al.*, 1997; Ahmed *et al.*, 2016).



Joonis 3. *CYP2D6* alleelide sagedused erinevates maailma populatsioonides (Gaedigk *et al.*, 2017). (Tõlgitud ja kohandatud).



Joonis 4. *CYP2D6* metabolismitüüpide keskmine esinemissagedus Euroopa ja Põhja-Ameerika populatsioonis (PM 5-10%; IM 10-17%; EM 70-80%; UM 3-5%) (Sachse *et al.*, 1997; Ahmed *et al.*, 2016).

1.3 Farmakogeneetilised uuringud

Viimase 10 aasta jooksul on toimunud DNA sekveneerimismeetodite vallas väga kiire areng ning kasutusel on väga kõrge läbilaskevõimega meetodeid, mille rakendus muutub järjest odavamaks. Tänu inimese referentsgenoomi avaldamisele on DNA sekveneerimistehnoloogiad hüppeliselt arenenud ning muutunud ka kättesaadavaks farmakogenoomilistes uuringutes. Teostatakse teise põlvkonna sekveneerimist (*next generation sequencing*, NGS) ning sellega on tehtud ka mitmeid ülegenoomseid assotsiatsiooniuuringuid (*genome-wide association studies*, GWAS) (Daly, 2010). Farmakogenoomikas kasutatakse GWAS uuringut, et identifitseerida genoomis aset leidvad ühe nukleotiidi (A, T, C või G) muutused, mis mõjutavad ravimivastust ning ADR-ide esinemist. Kui SNP-e esineb geenides, mis kodeerivad ensüüme, transportereid, ionkanalite komponente, mis on olulised ravimi toimes organismis, siis võib muutuda sellega ka ravimi farmakokineetika ja –dünaamika (Maggo *et al.*, 2016).

NGS-is kasutatakse kas kogu genoomi sekveneerimist (*whole genome sequencing*, WGS) või ainult eksoomi sekveneerimist (*whole exome sequencing*, WES). Viimane koondab enda alla kõik valke kodeerivad alad inimese DNA-st ehk eksonid. WES võimaldab kindlaks teha valke kodeerivates regioonides esinevaid variatsioone odavamalt ja lihtsamalt kui WGS (Maggo *et al.*, 2016). Paraku on selle meetodi miinuseks, et ei ole võimalik detekteerida variatsioone, mis asuvad väljaspool eksoneid, mistõttu ei saa ka analüüsida mittekodeerivate alade tähtsust. WGS aga seda võimaldab ning lisaks saab määrata ka CNV-sid, mis samuti mõjutavad haiguste teket ning ravimimetabolismi. Paraku on WGS hetkel veel küllaltki kallid uurimismeetodid (Meienberg *et al.*, 2016). Lisaks on kasutusel ka mitmeid genotüpiseerimise platvorme, mis farmakogeneetilistes analüüsides määravad SNP-e geenides, mis on olulised ravimi ADME-s (Arbitrio *et al.*, 2016). Farmakogeneetika ja farmakogenoomika esimesteks uurimismeetoditeks olid kandidaatgeenide (*candidate genes*) uurimine ning ravimimetabolismi ensüümi aktiivsuse fenotüpiseerimine. Kandidaatgeenide puhul uuritakse kindlaid ravimi metabolismi, transpordi või toksilisusega ning ka immuunvastuse tekkimisega seotud geene (Low *et al.*, 2014).

1.4 Farmakogeneetilistest uuringutest kinnitatud geen-ravim seosteni

1.4.1 PharmGKB

The Pharmacogenomics Knowledge Base (PharmGKB, www.pharmgkb.org) on 2000. aastal loodud kliiniliste farmakogenoomiliste (PGx) biomarkerite andmebaas, mis annab edasi olulist informatsiooni ravimite ja geenide koostoime kohta, hinnates juba seni avastatud geneetiliste varieeruvuste olulisust ravimi metabolismis. PharmGKB kogub, töötleb ning jagab andmeid geen-ravim toimete ning genotüüp-fenotüüp seoste kohta. PharmGKB-s saadav informatsioon on jagatud kuude erinevasse sektsiooni: variatsioonide annotatsioonid (*variant annotations*, VA), kliinilised annotatsioonid (*clinical annotations*, CA), ravimi metabolismi rajad (*pathways*), väga olulised farmakogeenid (*very important pharmacogene*, VIP), PGx põhiseid doseerimissuunised (*PGx prescribing information*) ning ravimi infolehed (*drug labels*) (www.pharmgkb.org/whatIsPharmgkb, külastatud 24. mai 2018).

Variatsioonide annotatsioonide sektsioonist leiab olulist informatsiooni geneetilise variatsiooni (SNP, CNV, deletsioon, insertsioon jt) ja ravimi fenotüüpide seoste kohta. Kuraatorid kombineerivad erinevate teaduslike publikatsioonide geen-ravim seosed ühtseks VA-ks, võttes

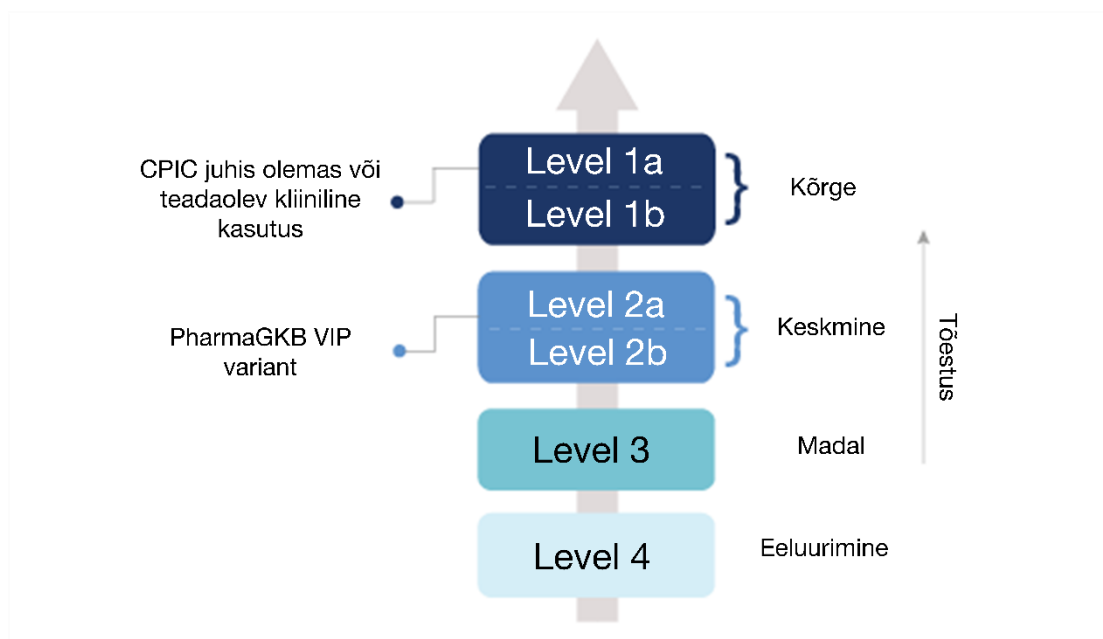
arvesse nii positiivseid kui ka negatiivseid geen-ravim seoseid. Andmebaasi kuraatorid järgivad kindlaid määratletud parameetreid (uuringugrupi suurus, etniline päritolu, alleelisageduste esinemine grupis, tulemuste statistiline olulisus), mille põhjal otsustatakse seose olulise üle. (McDonagh *et al.*, 2011).

Kliiniline annotatsioon tagab mitme erineva teaduslikult tõestatud VA põhjal kliinilise interpretatsiooni geen-ravim seoste kohta, mida arstid saaksid haiglates mugavalt kasutada. Igale CA-le määratakse n-ö tõestatuse tase (*level of evidence*). Taseme määramisel arvestatakse VA-de positiivsete ja negatiivsete seoste arvu, uurimisgrupi suurst ning statistilist olulisust (p-väärtus, šansisuhe). Tasemed on vastavalt kõrgemast madalamaks 1A, 1B, 2A, 2B, 3, 4 (Joonis 5) (Whirl-Carrillo *et al.*, 2012).

VIP-id on väga olulised farmakogeenid, mis on valitud kindlate kriteeriumite järgi. VIP-id on geenid, mis on olulised mitme ravimi metabolismis või on neis geenides kirjeldatud variatsioone, mis võivad tekitada ADR-e. VIP-ide annotatsioonist leiab informatsiooni geeni struktuuri kohta, seoseid haiguste esinemises ja kodeeriva valgu aktiivsuse mõjutamises, sealhulgas mõju ravimi farmakokineetikale ja –dünaamikale. Hetkel on PharmGKB andmebaasis määratud 65 VIP geeni (www.pharmgkb.org/vips, külastatud 24. mai 2018), (McDonagh *et al.*, 2011).

Samuti leiab PharmGKB andmebaasist jooniseid ja diagramme oluliste geenide metabolismiradade kohta, mis näitavad konkreetse toimeaine interaktsioone farmakogeenide poolt kodeeritud valkudega. Viimaks koondab PharmGKB ka suunised ravimite väljakirjutamiseks. Sisestades kindla geeni genotüübi informatsiooni andmebaasi, on võimalik indiviidil saada ülevaade endale sobivaimast ravimiannusest. Doseerimissuuniste annotatsioonid on kirjutanud CPIC, *Dutch Pharmacogenetics Working Group* (DPWG), *Canadian Pharmacogenomics Network for Drug Safety* (CPNDS) jt ekspertorganisatsioonid (McDonagh *et al.*, 2011).

PharmGKB on ka koondanud informatsiooni teadustöodes tuvastatud ja farmakogeneetikas oluliste geenide haplotüüpidest. Andmebaas jagab definitsioonitabeleid, mis peaksid lihtsustama geneetiliste variatsioonide põhjal haplotüüpide, star-alleelide määramist. Lisaks on olemas ka vastavuse tabelid, mis aitavad jõuda star-alleelide määramisest konkreetsete metabolismi fenotüüpideni (www.pharmgkb.org/page/pgxGeneRef, külastatud 24. mai 2018).



Joonis 5. PharmGKB tõestatuse tasemed geen-ravim paaride kohta

(www.pharmgkb.org/page/clinAnnLevels, külastatud 24. mai 2018) (Tõlgitud ja kohandatud). Hetkel on tõestatuse tase määratud kõrgemalt tasemelt madalamale liikudes vastavalt 46, 10, 70, 90, 59 ja 6 geen-ravim paarile (www.cpicpgx.org/genes-drugs/, külastatud 27. mai 2018).

1.4.2 CPIC

CPIC (*Clinical Pharmacogenetics Implementation Consortium*) on 2009. aastal rajatud rahvusvaheline konsortsium, kuhu kuuluvad vabatahtlikud farmakogenoomika, -geneetika ja laboratoorse meditsiini eksperdid. Nende eesmärgiks on suurendada farmakogeneetiliste testide kättesaadavust igapäevases kliinilises meditsiinis. CPIC pakub kontrollitud, uuendatud, teadsulikult tõestatud ning kõigile kättesaadavat informatsiooni geenide ja ravimite koostoimest (Relling ja Klein, 2011).

CPIC juhtnööride eesmärgiks on aidata arstidel mõista geneetiliste testide tulemusi ning kuidas nende põhjal määrata patsiendile efektiivne ja ohutu ravim. Genotüüp-fenotüüp seose kohta otsitakse kinnitust ilmunud teadustöödest ja –artiklitest. Igale genotüübi ja fenotüübi paarile antakse hinnang skaalal „kõrge“, „mõõdukas“ ja „nõrk“. „Kõrge“ puhul on läbi viidud põhjalikud uurimistööd ning tulemused on olnud kooskõlas teiste ilmunud teadustöödega. „Mõõduka“ korral on piisavalt tõendeid, kuid nende tugevus on kaheldav väheste individuaalsete uuringute tõttu.

„Nõrga“ hinnangu puhul ei ole piisavalt tõendeid ja/või teadustöodes esineb vigu (www.cpicpgx.org/levels-of-evidence/, külastatud 27. mai 2018). Samuti antakse ravimi väljakirjutussoovitusi vastavalt geen-ravim paarile (CPIC level A-D). Level A geen-ravim paari puhul on leitud piisavalt tõendeid ning tuleks kasutada alternatiivseid ravimeid või muuta ravimiannust. Level D ei soovita muuta ravimi väljakirjutamist, sest puuduvad piisavad tõendid ning katsed geen-ravim toime kohta (www.cpicpgx.org/prioritization/#cpicLevels, külastatud 27. mai).

Seega nii CPIC kui ka PharmGKB andmebaas sisaldavad uuringute tulemustena leitud põhjalikku infot ravimivastustes oluliste geneetiliste alleelide ning nende määramise kohta. Niisamuti esitletakse andmebaasides alleelide seoseid ravimitega ning neid saaks juba kasutada konkreetse populatsiooni andmete analüüsiks ning edaspidiseks personaalsemaks lähenemiseks ravimiteraapias.

2. UURIMUS

2.1 Töö eesmärk

Antud bakalaureusetöö esimeseks eesmärgiks on Tartu Ülikooli Eesti Geenivaramu (TÜ EGV) andmete põhjal analüüsida ravimivastuses olulise tsütokroom P450 perekonna viie erineva geeni alleelide esinemissagedusi ja iseloomustada ka nende põhjal ennustatud fenotüüpe Eesti populatsioonis ning võrrelda tulemusi PharmGKB andmebaasi europiidse populatsiooni alleelide sagedustega. Teiseks, on eesmärgiks hinnata TÜ EGV täisgenoomi andmetes geeni *CYP2D6* koopiaarvude (CNV) esinemist arvutuslikul teel ja võrrelda esinemissagedusi teise sõltumatu meetodi tulemustega. Kolmandaks, on eesmärgiks valideerida arvutuslikult saadud mitte kokku langevad CNV-de määramised, kasutades *TaqMan* süsteemi.

2.2 Materjalid ja metoodika

2.2.1 Star-alleelide määramine

Antud töös kasutati viie *CYP* geeni alleelide sageduste määramiseks TÜ EGV-s juba olemasolevaid 2420 geenidoonori täisgenoomi andmeid (Tabel 2). Valim on koostatud nii, et see annaks võimalikult täpse ülevaate Eesti populatsioonist. Igale indiviidile määrati PharmGKB poolt jagatavate star-alleelide defineerimise tabelite alusel vastava viie geeni star-alleelide haplotüübid. Seejärel määrati võimalusel igale indiviidile ka star-alleeli diplotüübid ning leiti nende sagedused valimis. Saadud tulemusi võrreldi PharmGKB andmebaasis dokumenteeritavate *CYP* geenide diplotüüpide sagedustega europiidses populatsioonis. Lisaks leiti igale indiviidile diplotüübile vastav fenotüüp, mille määramine põhines samuti PharmGKB-s kajastatud informatsioonil. Põhiline andmete analüüs teostati MS Excelis.

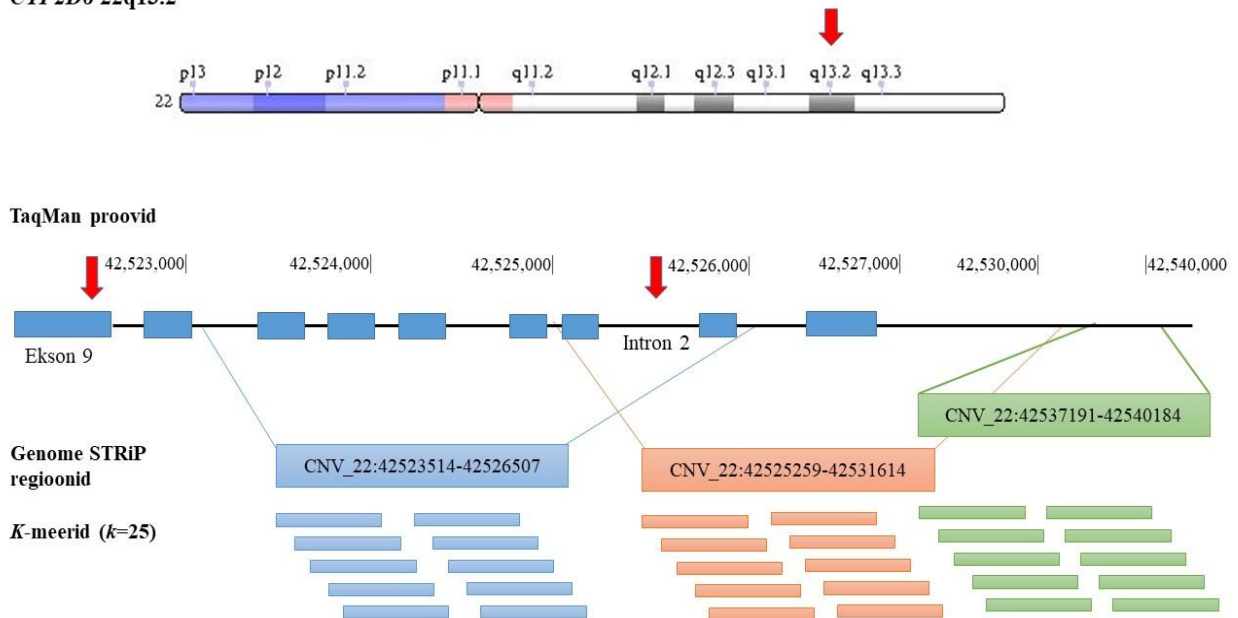
Tabel 2. Antud töös uuritud viis CYP geeni; geenide erinevate star-alleelide arv PharmGKB andmebaasis ja Eesti populatsioonis; ravimid, mille metabolismi mõjutavad. Andmed võetud PharmGKB, CPIC andmebaasist ja TÜ EGV-st.

Geen	Erinevate star-alleelide arv	Eesti populatsioonis leitud star-alleelide arv	Level A CPIC suunisega väljakirjutatavad ravimid
<i>CYP2C9</i>	60	8	Epilepsiaravim (fenütoiin); vere hüübimist takistav (varfariin)
<i>CYP2C19</i>	34	6	Antidepressandid (amitriptüliin, tsitalopraam, estsitalopraam); trombotsüüte pärssiv (klopidogreel); seenevastane ravim (vorikonasool)
<i>CYP2D6</i>	161	15	Antidepressandid (amitriptüliin, fluvoksamiin, nortriptüliin, paroksetiin); opioidid (kodeiin, oksükodoon); rinnanäärmevähi ravim (tamoksifeen); kemoteraapia järgne ravim (odansetroon, tropisetron)
<i>CYP3A5</i>	9	3	Immuunsupressiivne ravim (takroliimus)
<i>CYP4F2</i>	3	3	Vere hüübimist takistav (varfariin)

2.2.2 *CYP2D6* koopiaarvud

CYP2D6 koopiaarvud määrati TÜ EGV täisgenoomide andmetest kahe erineva arvutusliku meetodiga. Esimese meetodi puhul kasutati *Genome STRiP CNV Discovery Pipeline* (versioon 2.00.1611), et leida ja genotüpiseerida suuri deletsioone, duplikatsioone ning multialleelseid CNV-sid sekveneeritud täisgenoomidest. CNV-d määratakse populatsiooni WGS andmetest, kasutades programme, mis lugemite katvuse põhjal otsivad genoomis variatsioone (Handsaker *et al.*, 2015). Leitud geeni piirkonnad genotüpiseeriti kasutades *Genome STRiP SVGenotyper* moodulit. Tulemuste võrdlemiseks kasutati teist meetodit, mis kasutas *k*-meeride nimekirja ($k=25$). Üheastmelised 25-meerid tehti kasutades *GenomeTester4* paketi *GListMaker* programmi (Kaplinski *et al.*, 2015). Esimene ja teine meetod määrasid koopiaarve kolmes erinevas *CYP2D6* geeni regioonis, vastavalt 22:42523514-42526507, 22:42525259-42531614 ja 22:42537191-42540184 (Joonis 6). Esimeses meetodis uuriti 2269 täisgenoomi ning teises 1873. Töös leiti *CYP2D6* regioonide CNV sagedused mõlema arvutusliku meetodi jaoks eraldi ning võrreldi neid omavahel. Arvutusliku meetodi tulemuste põhjal valiti välja 230 indiviidi, kellel oli *k*-meer meetodi järgi 2-st erinev koopiaarv ning valideerimiseks analüüsiti *CYP2D6* geeni koopiaarvude esinemist kahe erineva *TaqMan* koopiaarvu analüüsiga (*TaqMan Copy Number Assay*). Esimeses analüüsis oli sihtmärk regiooniks *CYP2D6* geeni ekson 9 (analüüsi id: Hs00010001_cn) ja teises intron 2 (analüüsi id: Hs04083572_cn) (Joonis 6). *TaqMan*'iga ja arvutuslikult tuvastatud koopiaarvude sagedusi võrreldi ning leiti korrelatsioon kahe meetodi vahel. Põhiline andmete analüüs teostati MS Excelis.

CYP2D6 22q13.2



Joonis 6. CYP2D6 asukoht kromosoomis. *Genome STRiP* meetodis kasutatud geeni piirkonnad. *TaqMan* ekson 9 ja intron 2 analüüsi regioonid ja 25-meerid (Kromosoomi joonis võetud www.ghr.nlm.nih.gov/gene/CYP2D6#location, külastatud 24. mai 2018; lookuse ülesehitus www.genome.ucsc.edu põhjal, külastatud 24. mai 2018; *TaqMan* proovide joonis tehtud *ThermoFisher Scientific*'u CYP2D6 genoomikaardi põhjal, www.thermofisher.com/order/genome-database/details/copy-number/Hs00010001_cn, külastatud 24. mai 2018)

2.3 Tulemused

2.3.1 Star-alleelide määramine

Bakalaureusetöös analüüsiti 5 *CYP* geeni allelele, mis kodeerivad olulisi ensüüme faas I ravimi metabolismi reaktsioonides (Zanger ja Schwab, 2013; Zhou *et al.*, 2017). Teostatud analüüs näitab, et ligikaudu 47% sagedusega esineb geenidoonoritel tavapärasest metabolismist erinev fenotüüp vähemalt ühes geenis viiest uuritud *CYP* geenist. Antud valimis tuvastati *CYP2C9* geenil 16 erinevat diplotüüpi (Tabel 3). Domineerivaks diplotüübiks TÜ EGV andmetes on $*1/*1$ (70,4%), mida oli ka oodata, kuna tegemist on referentsalleeliga ehk antud diplotüübile vastav fenotüüp on *CYP2C9* tavapärane metaboliseerija. Sagedasemad diplotüübid on veel $*1/*2$ ja $*1/*3$, mida esineb vastavalt 13,9% ning 11,9% geenidoonoritel. Mõlemad diplotüübid määravad fenotüübiks IM (www.pharmgkb.org/page/cyp2c9RefMaterials, külastatud 27. mai 2018). 2420 geenidoonori

WGS andmete põhjal on alleelisagedused sarnased Euroopa alleelisagedustega, vastavalt *1/*1 (64%), *1/*2 (20,1%), *1/*3 (11,3%). (Tabel 3).

Tabel 3. *CYP2C9* geeni diplotüübid ja nende esinemissagedus TÜ EGV doonorite hulgas, euroopiidses populatsioonis ning vastav fenotüüp. Andmed võetud TÜ EGV täisgenoomidest ja PharmGKB andmebaasist (n=2420).

Diplotüüp	Indiviidide arv	Esinemissagedus geenidoonoritel	PharmGKB sagedused	Fenotüüp
*1/*1	1703	70,37%	64,01%	NM
*1/*2	337	13,93%	20,17%	IM
*1/*3	287	11,86%	11,33%-	IM
*2/*3	37	1,53%	1,79%	PM
*1/*29	15	0,62%	-	Määramata
*2/*2	14	0,58%	1,59%	PM
*1/*12	8	0,33%	-	IM (arvatav)
*3/*3	7	0,29%	0,50%	PM
*3/*29	3	0,12%	-	Määramata
*1/*11	2	0,08%	0,27%	IM (arvatav)
*2/*29	2	0,08%	-	Määramata
*1/*4	1	0,04%	-	IM (arvatav)
*1/*45	1	0,04%	-	Määramata
*12/*12	1	0,04%	-	PM (arvatav)
*12/*29	1	0,04%	-	Määramata
*3/*11	1	0,04%	0,02%	PM (arvatav)

PharmGKB andmetel on leitud *CYP2C19* geenil 12 erinevat star-alleeli europiidses populatsioonis. TÕ EGV andmetest tuvastati 6 star-alleeli, mis moodustasid 11 erinevat diplotüüpi (Tabel 4). Diplotüübid *1/*1 ja *1/*17 on kõige enam esinevad, vastavalt 35,9% ja 31,2%. Alleelide *1/*1 kandja on tavapärase fenotüübiga, kuid *1/*17 kandja on tavapärasest kiirema fenotüübiga. Euroopa ja Põhja-Ameerika populatsioonis domineerivad samuti diplotüübid *1/*1 (39%) ja *1/*17 (13,3%), vastavalt PharmGKB andmebaasile. See-eest võrreldes PharmGKB andmetega, on diplotüübi *1/*17 sagedus TÕ EGV geenidonorite seas umbes kaks korda suurem kui europiidses populatsioonis. (Tabel 4).

Tabel 4. *CYP2C19* geeni diplotüübid ja nende esinemissagedus TÕ EGV doonorite hulgas, europiidses populatsioonis ning vastav fenotüüp. Andmed võetud TÕ EGV täisgenoomidest ja PharmGKB andmebaasist (n=2420).

Diplotüüp	Indiviidide arv	Esinemissagedus geenidonoritel	PharmGKB sagedused	Fenotüüp
*1/*1	868	35,87%	39,00%	NM
*1/*17	755	31,20%	13,30%	RM
*1/*2	368	15,21%	9,10%	IM
*2/*17	188	7,77%	3,11%	IM
*17/*17	166	6,86%	4,54%	UM
*2/*2	57	2,36%	2,12%	PM
*1/*4A	8	0,33%	0,17%	IM
*1/*8	4	0,17%	0,16%	IM
*1/*3	3	0,12%	0,39%	IM
*2/*4A	2	0,08%	0,04%	PM
*8/*17	1	0,04%	0,05%	IM

CYP2D6 geeni lookus on väga polümorfne ja keerukas, seetõttu on geenil avastatud 119 erinevat alleeli, millest 22-l on PharmGKB andmebaasis sagedused europiidses populatsioonis veel määramata. Kuna *CYP2D6* star-alleelide määramiseks on vaja genotüpiseerida umbes 150 erinevat SNP-i, võib osutuda star-alleelide üheselt tuvastamine keerukaks. TÕ EGV täisgenoomi andmetest leiti 29 erinevat diplotüüpi (Tabel 5), kusjuures 1337 (55,3%) geenidoonoril ei olnud võimalik mõlemat alleeli määrata ning 921-l (38,1%) tuvastati ainult üks alleel kahest. Määramatu viitab sellele, et ei leitud vastavat järjestust SNP-de tabelist või järjestus kattus mitme star-alleeliga, mistõttu oli raske määrata üht kindlat alleeli. Indiviididest, kellel õnnestus määrata kaks alleeli (162), on 79-l (3,26%) diplotüüp *4/*4, mis määrab PM fenotüübi. Europiidses populatsioonis on *4 alleeli sagedus 18,2% ning diplotüübi *4/*4 sagedus 3,3%, vastavalt PharmGKB andmetele. TÕ EGV täisgenoomi andmetest <1% leidub ka haruldasi *CYP2D6* allelele ja diplotüüpe, mille sagedusi ning ennustatud fenotüüpe europiidses populatsioonis ei ole veel PharmGKB andmete põhjal määratud. Näiteks, *CYP2D6* alleelid *22, *83, *108 (Tabel 5).

Tabel 5. Tuvastatud *CYP2D6* geeni diplotüübid ning vastav fenotüüp. Esinemissagedus TÜ EGV doonorite hulgas ning võrdlus europaalse populatsiooniga. Andmed võetud TÜ EGV täisgenoomidest (n=2420) ja PharmGKB andmebaasist.

Diplotüüp	Indiviidide arv	Esinemis- sagedus geenidoonoritel	PharmGKB sagedused	Fenotüüp
*4/*4	79	3,26%	3,30%	PM
*3/*4	20	0,83%	0,49%	PM
*4/*33	17	0,70%	0,69%	NM
*4/*9	13	0,54%	0,72%	IM
*3/*33	5	0,21%	0,05%	NM
*33/*33	4	0,17%	0,04%	NM
*4/*22	4	0,17%	Määramata	Määramata
*3/*3	3	0,12%	0,02%	PM
*4/*6	3	0,12%	0,35%	PM
*9/*33	2	0,08%	Määramata	NM
*4/*83	1	0,04%	Määramata	Määramata
*3/*22	1	0,04%	Määramata	Määramata
*6/*6	1	0,04%	0,009%	PM
*3/*9	1	0,04%	0,05%	IM
*6/*9	2	0,08%	0,04%	IM
*4/*108	1	0,04%	Määramata	Määramata
*9/*22	1	0,04%	Määramata	Määramata
*22/*22	1	0,04%	Määramata	Määramata
*6/*22	1	0,04%	Määramata	Määramata
*3/*6	1	0,04%	0,03%	PM
*9/*9	1	0,04%	0,04%	NM
Määramatu	2258	93,3%	-	-

Geenis *CYP3A5* leiti TÜ EGV täisgenoomi andmetest 4 erinevat diplotüüpi (Tabel 6). Kõige enam ehk 85,8% geenidoonoritest on diplotüüp *3/*3, millele vastab fenotüüp PM.. Diplotüüpe *1/*3 (IM), *1/*1 ja *3/*6 (PM) esineb geenidoonorite seas vastavalt 13,5%, 0,6% ning 0,04%. Euroopa ja Põhja-Ameerika populatsioonide diplotüüpide sagedused on sarnased TÜ EGV andmetega, vastavalt 84,8%, 14,4%, 0,6%, 0,18%. (Tabel 6).

Tabel 6. Tuvastatud *CYP3A5* geeni diplotüübid ning vastav fenotüüp. Esinemissagedus TÜ EGV doonorite hulgas ning võrdlus euroopiidse populatsiooniga. Andmed võetud TÜ EGV täisgenoomidest (n=2420) ja PharmGKB andmebaasist.

Diplo tüüp	Indiviidide arv	Esinemissagedus geenidoonoritel	PharmGKB sagedused	Fenotüüp
*3/*3	2077	85,83%	84,82%	PM
*1/*3	327	13,51%	14,37%	IM
*1/*1	15	0,62%	0,61%	NM
*3/*6	1	0,04%	0,18%	PM

CYP4F2 geenil leiti TÜ EGV täisgenoomide andmetest 5 erinevat diplotüüpi, millel oli võimalik määrata kahte alleeli (Tabel 7). Geenidoonorite seas esines 46 (1,9%) indiviidil määramatu diplotüüp. Nendest 26,3% suudeti tuvastada üks alleel, kuid 28,2% ei suudetud tuvastada kumbagi alleeli. *CYP4F2* diplotüüp *1/*1, mis viitab tavapärasele metabolismile, leiti 1525 (63%) indiviidil geenidoonoritest. 200 (8,3%) indiviidil määrati diplotüübiks *1/*3. Diplotüüpide *1/*2, *2/*3, *3/*3 sagedused jäävad alla 1% (Tabel 7). Euroopa ja Põhja-Ameerika populatsioonides on leitud sagedus ainult alleelile *CYP4F2**3 ning selle põhjal arvutati sagedus diplotüübile *3/*3, vastavalt 29,9% ja 8,9%. PharmGKB *CYP4F2* alleeli definitsiooni tabelis on fenotüüpi kirjeldatud ainult alleelidele *1, *2 ja *3, vastavalt tavapärane, tavapärane või suurenenud ning vähenenud ensüümi aktiivsus.

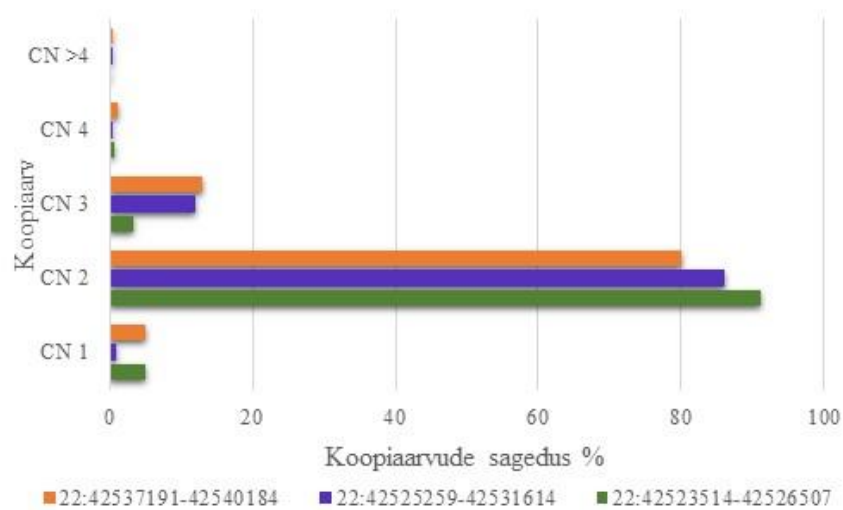
Tabel 7. Tuvastatud *CYP4F2* geeni diplotüübid ning vastav fenotüüp. Esinemissagedus TÜ EGV doonorite hulgas ning võrdlus Euroopa populatsiooniga. Andmed võetud TÜ EGV täisgenoomidest (n=2420) ja PharmGKB andmebaasist.

Diplo tüüp	Indiviidide arv	Esinemissagedus geenidoonoritel	PharmGKB sagedused	Fenotüüp
*1/*1	1525	63,02%	-	NM
*1/*3	200	8,26%	-	-
*3/*3	7	0,29%	8,98%	Vähenenud aktiivsus
*2/*3	4	0,17%	-	-
*1/*2	1	0,04%	-	-
Määramatu	683	28,22%	-	-

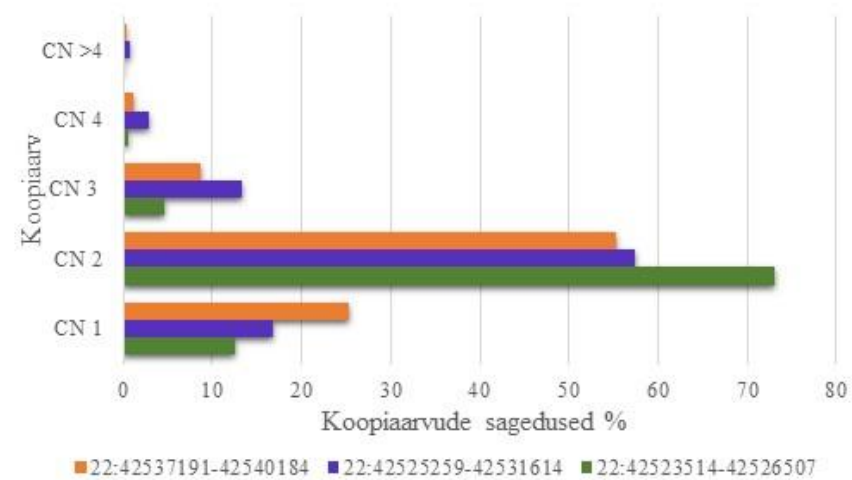
2.3.2 *CYP2D6* koopiaarvud

Analüüsidest TÜ EGV 2269 geenidoonori *CYP2D6* geeni koopiaarvude esinemist kolme erineva regiooni alusel, järeldeb esimese arvutusliku *Genome STRiP* meetodi põhjal, et regiooni 22:42523514-42526507 esineb 0-5 koopias. 2068 (91,1%) geenidoonoril on vastav piirkond geenis kahes koopias, 110 (4,9%) ühes koopias, 76 (3,4%) kolmes koopias ning vähem kui 1% doonoritest esineb koopiaarv null (0,04%), neli (0,6%) ja viis (0,04%) (Joonis 7A). Regioonis 22:42525259-42531614 varieeruvad koopiaarvud 0-7 vahel. Kõige rohkem on koopiaarvu 2, mis esineb 1816 ehk 80% geenidoonoritest. Koopiaarvu 3 on 304 indiviidil (13%), 1 on 109-l (5%) ning 4 on 28 (1,2%) doonoril. Koopiaarve 0, 5, 6 ja 7 esineb <1% (Joonis 7A). Kolmandas regioonis, 22:42537191-42540184, varieeruvad koopiaarvud 1-7 vahel, kuid seda regiooni ei leitud ühelgi geenidoonoril viies koopias. 1945 (86%) geenidoonoril on koopiaarvuks kaks, 272-l (12%) kolm, 33-l (1%) on üks. Koopiaarvude 4, 6, 7 sagedused on vastavalt 0,4%, 0,4% ning 0,1% (Joonis 7A). Kokkuvõtvalt, *Genome STRiP* meetodiga leiti kõige sagedasemaks koopiaarvuks geenidoonoritel kaks kõigis kolmes *CYP2D6* geeni regioonis.

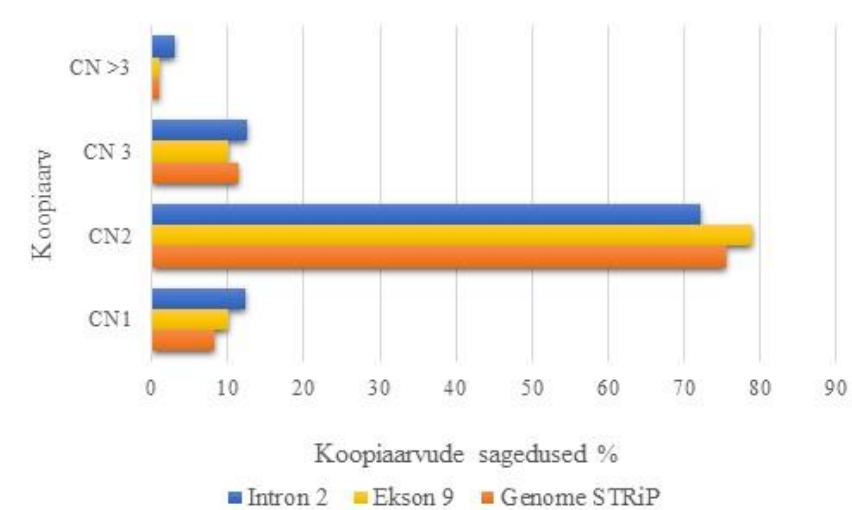
A



B



C



Joonis 7. Leitud koopiaarvud CYP2D6 geeni kolmele regioonile (22:42523514-42526507, 22:42525259-42531614, 22:42537191-42540184) TÜ EGV täisgenoomi andmetest. A) Arvutusliku *Genome STRiP* meetodiga määratud koopiaarvude sagedused (n=2269). **B)** *K*-meer meetodiga arvutuslikult leitud koopiaarvude sagedused (n=1873). **C)** *TaqMan Copy Number Assay*’ga määratud koopiaarvude esinemissagedused *Genome STRiP* arvutusliku meetodiga võrreldes (n=230).

Võrdlemiseks kasutati teist arvutuslikku meetodit, kasutades *GListMaker* programmiga tehtud *k*-meere. Käesoleva töö teostamise perioodil oli võimalik leida 2269 geenidonorist 1873-le koopiaarvud *k*-meer meetodil samuti kolmele eelnevalt mainitud *CYP2D6* geeni regioonile. Sarnaselt eelmisele arvutuslikule meetodile leiti analüüsis domineerivaimaks koopiaarvuks kaks. Esimene regioon, 22:42523514-42526507, on 1369 (73,1%) geenidonoril kahes koopias, 236-l (12,6%) on ühes, 88-l (4,7%) on kolmes ning 10 (0,5%) indiviidil on antud regioon neljas koopias. Regiooni 22:42525259-42531614 tuvastati 0-10 koopias. Geenidonoritest 1074-l (57,3%) leiti koopiaarvuks kaks, 315-l (16,8%) kolm, 248-l (13,2%) üks ning 52-l (2,8%) leiti neli. Koopiaarvud, mis on suuremad kui neli, esinevad <1% geenidonoritest. Kolmandat regiooni, 22:42537191-42540184, on valdavalt kahes koopias, vastavalt 1035 (55,3%) geenidonoril 1873-st. Järgnevad 474 (25,3%) geenidonorit ühe koopia, 163 (8,7%) kolme koopia ning 22 (1,2%) nelja koopiaga. Vähem kui 1% geenidonoritest leiti koopiaarvuks 5, 6 või 8 (Joonis 7B).

Võrreldes omavahel *Genome STRiP* ja *k*-meer meetodiga (n=1873) leitud koopiaarve selgub, et kolme *CYP2D6* regiooni alusel leitud koopiaarvud langevad kokku keskmiselt 66% ulatuses kahe arvutusliku meetodi vahel, vastavalt 78,1%, 64,3%, 57,5%. Esimeses regioonis erinesid koopiaarvud 230 indiviidil, teises regioonis 489-l ja kolmandas geeni regioonis 616 geenidonoril. Esimeses regioonis tuvastas *Genome STRiP* meetod kõige rohkem koopiaarvu kaks, 201/230-st. *K*-meer meetodiga leiti aga koopiaarvuks enamasti 1 (155 indiviidil) ning mõnel juhul ka 2-st suurem koopiaarv. Lisaks oli valimis 180 indiviidi, kellel oli koopiaarv määramata, kas *Genome STRiP* või *k*-meer meetodiga, nii et tulemusi võrrelda ei saanud.

Arvutuslike meetodite võrdlemisel võeti välja 230 proovi ning valideerimiseks kasutati kahte *TaqMan* koopiaarvu analüüsi. Esimesel juhul leiti koopiaarvud geeni *CYP2D6* ekson 9 regiooni alusel, teisel juhul intron 2. *TaqMan* meetoditega ja *Genome STRiP*’iga määratud koopiaarvude tulemused 230 indiviidile on esitatud joonisel 7C. Nii nagu arvutuslike meetoditega, leiti ka

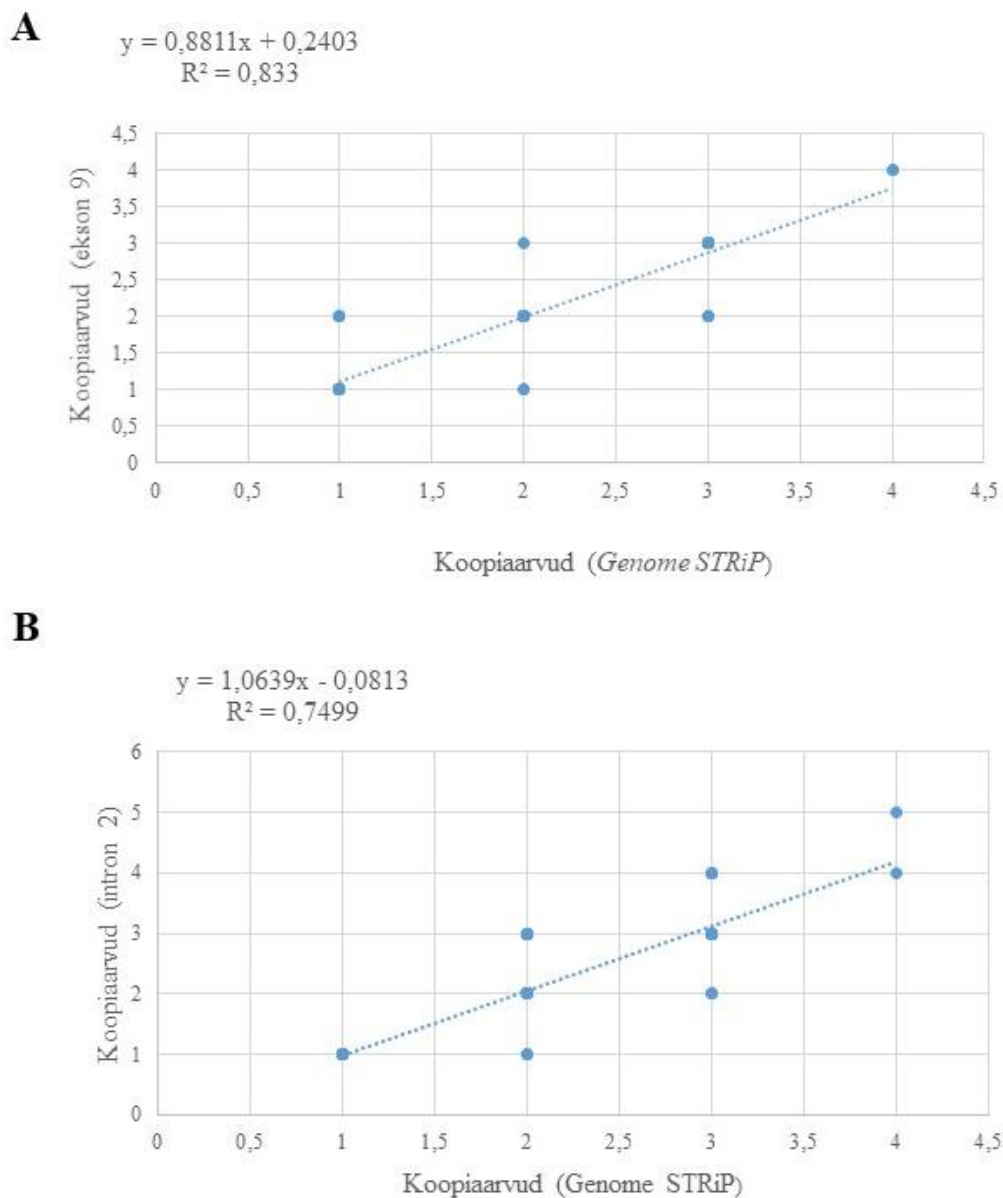
TaqMan analüüsiga kõige sagedasemaks koopiaarvuks geenidonoritel kaks (Joonis 7C). *TaqMan*'iga valideeritud 230-st proovist 71-l olid koopiaarvud määratud nii *Genome STRiP* kui ka *k*-meer meetodiga. Kui võtta välja individid, kelle koopiaarvud erinesid kahe arvutusliku meetodi vahel (12/71-st), siis selgub, et *TaqMan* ekson 9 ja intron 2 alusel määratud koopiaarvud langevad kokku *Genome STRiP* meetodi tulemustega. Valimist 48 geenidoonori leitud koopiaarvud klappisid nii *Genome STRiP*, *k*-meer kui ka *TaqMan* ekson 9 ja intron 2 analüüsiga.

Kahe arvutusliku meetodi omavahelisel võrdlemisel (n=1873) ei klappinud 230 proovi koopiaarvud. Nendest 230-st oli *TaqMan* analüüsi valimis paraku ainult 17 indiviidi. Ekson 9 analüüs määras neile 17 geenidoonorile samad koopiaarvud, mis *Genome STRiP* ning intron 2'ga leiti 14 indiviidile sama koopiaarv, mis *Genome STRiP*'ga (Lisa 1).

Selleks, et hinnata eelnevate tulemuste kokkulangevust täpsemalt, leiti korrelatsioon antud töös uuritud esimese arvutusliku *Genome STRiP* meetodi esimeses regiooni, 22:42523514-42526507, ja kahe *TaqMan* analüüsi tulemuste vahel. Mõlemal juhul leiti tugev positiivne lineaarne seos ($r^2=0,833$ ja $r^2=0,7499$; $p<0,05$) (Joonis 8). Kõigist *TaqMan* 'iga valideeritud täisgenoomidest (n=230) klappisid 199 geenidoonori *CYP2D6* geeni koopiaarvude andmed kolme erineva meetodi vahel. Lisaks olid määratud ka 96 indiviidile mõlemad *CYP2D6* geeni alleelid (Lisa 2). *Genome STRiP*'i ja ekson 9 analüüsi võrdluses selgus, et 10 indiviidil 230-st erinesid leitud koopiaarvud (Lisa 2). Suurimaks erinevuseks kahe meetodi vahel oli arvutusliku meetodiga leitud geeni deletsiooni (koopiaarv 1) asendumine funktsionaalse geeni koopiaga (koopiaarv 2) ekson 9 analüüsis. Samas, intron 2 *TaqMan* analüüsiga leiti nendest 10-st indiviidist 8-le sama koopiaarv, mis saadi arvutusliku meetodiga (Lisa 3). Nendest kümnest proovist, mille tulemused omavahel ei kattunud, leiti *k*-meer meetodiga samas piirkonnas koopiaarvud ainult kolmele indiviidile. Ühel indiviidil kattus *k*-meeri koopiaarv *Genome STRiP*'ga, ühel kattus ekson 9 meetodiga ning kolmandal erines mõlemast eelnevast meetodist.

Intron 2 *TaqMan* analüüsiga oli koopiaarvu erinevusi võrreldes arvutusliku *Genome STRiP* meetodiga 23 proovil 230-st (Lisa 4). Kui *Genome STRiP* meetodiga leiti geenidonoritel *CYP2D6* geenile koopiaarvuks 2, mis tagab tavapärase metabolismi, siis intron 2 analüüsis saadi koopiaarvuks 3 ehk geeni duplikatsioon, mis viitab kiiremale metabolismile. Nendest 23 proovist 21 *Genome STRiP*'i tulemused ühtisid ka ekson 9 *TaqMan* analüüsi tulemustega (Lisa 3). *K*-meer meetodiga leiti koopiaarve 11 indiviidile 23-st, kellele määratud koopiaarvud ei klappinud *Genome STRiP*'i ja intron 2 meetodite vahel. Viis leitud koopiaarvu ühtisid *Genome STRiP*'i tulemustega,

kaks klappisid intron 2'ga ning 4 indiviidi koopiaarvud erinesid kõigist. Kahe *TaqMan* meetodi omavaheline võrdlus annab koopiaarvude kokkulangevuse 200 proovil 230-st, 30 indiviidi koopiaarvud erinevad.



Joonis 8. Koopiaarvude korreltasioon. A) *Genome STRiP*'i esimese regiooni, 22:42523514-42526507, ja ekson 9 *TaqMan* meetodi vahel (n=230). B) *Genome STRiP*'i esimese regiooni, 22:42523514-42526507, ja intron 2 *TaqMan* meetodi vahel (n=230).

2.4 Arutelu

Ravimi metabolismi mõjutavad mitmed tegurid, kuid peamine neist on geneetiline polümorfism. Analüüsides patsientide ravimi metabolismi geene, on suure tõenäosusega võimalik ennustada neist igatüüpi ravimitoime efektiivsust. Tänu CPIC ja PharmGKB andmebaasidele on tänapäeval võimalik geeni alleelide põhjal määrata konkreetse ravimi metabolismi fenotüüpi erinevates populatsioonides ja etnilistes gruppides. Käesolevas bakalaureusetöös uuriti ravimi metabolismi *CYP* geeniperekonna diplotüüpide ning ennustatavate fenotüüpide sagedust Eesti populatsioonis ning hinnati *CYP2D6* koopiaarvude esinemissagedust erinevate meetoditega.

Käesolevas töös teostatud analüüsi tulemusena selgus, et *CYP* perekonna geenide *CYP2C9*, *CYP2D6*, *CYP3A5* diplotüüpide sagedused TÜ EGV täisgenoomide andmetes on sarnase sagedusega ka europiidses populatsioonis, tuginedes PharmGKB andmebaasi andmetele. Siiski saab erinevusena välja tuua geeni *CYP2C19*, mille diplotüübi **1/*17* sagedus Eesti populatsioonis on umbes kaks korda suurem kui europiidses populatsioonis PharmGKB andmete põhjal (www.pharmgkb.org/page/cyp2c19RefMaterials, *CYP2C19 Frequency Table*, külastatud 27. mai 2018). Erinevad uurimistööd on näidanud, et Euroopas ja Põhja-Ameerikas on keskmine diplotüübi sagedus <20% (Shirasaka *et al.*, 2016). Vaatamata sellele on Norras, Saksamaal ja Venemaal tuvastatud diplotüübi **1/*17* sagedused üsna sarnased antud bakalaureusetöös leitud sagedusega Eesti populatsioonis, vastavalt 26,5%, 36%, 39,8%. Norras kasutati uuringus terveid indiviide, Saksamaal olid fookuses südame pärgarteri haigustega patsiendid, ning Venemaal uuriti maohaavanditega patsiente (Pedersen *et al.*, 2010; Sibbing *et al.*, 2010; Sychev *et al.*, 2015). *CYP2C19* on oluline ensüüm trombotsüütide agregatsiooni pärssiva ravimi klopidooreel metabolismis. Klopidooreel on eelravimi kujul ning seega peab toimimiseks läbima biotransformatsiooni, milles on oluline roll *CYP2C19*-l (Arbitrio *et al.*, 2016; Scott *et al.*, 2013). Mitmed teadustööd kinnitavad, et *CYP2C19*17* pärssib klopidooreeli mõju ning võib tekitada sellega veritsustüsistusi ahenenud pärgarteritega patsientidel, kellele on tehtud perkutaanne koronaarinterventsioon (*percutaneous coronary intervention* PCI) ja kes on saanud tromboosi ennetavat ravi (Sibbing *et al.*, 2010). Teisalt aga on ka teadustöid, kus on leitud, et **17* alleel ei vähenda klopidooreeli tromboosi pärssivat efekti nii palju, et patsiendid, kes seda alleeli kannavad, peaksid ravi vahetama. Sellegipoolest on antud seost hinnatud CPIC poolt kui kõige kõrgema tõestatuse tasemega, level A, geen-ravim paariks (Joonis 5). See-eest indiviididel, kes kannavad **2* alleeli ehk vähenenud funktsiooniga fenotüüpi, esineb rohkem trombotsüütide agregatsiooni ka

pärast ravi ning soovitatud on kasutada alternatiivseid ravimeid (Scott *et al.*, 2013; Shirasaka *et al.*, 2016). Antud alleel esines käesoleva bakalaureusetöö andmetel 25,4% geenidoonoritel. Sellest tulenevalt on oluline tulevikus edasi uurida, kas TÜ EGV geenidoonoritel, kelle tuvastati *CYP2C19* geenis *17 alleel või *2 alleel, on diagnoositud pärgarterite haigusi ning millist ravi neile on määratud, et vältida ohtlikke kõrvalmõjusid. Samamoodi tuleks geenidoonoritele tagasiside andmisel informeerida ka kõiki neid inimesi, kellel esineb alleel, aga ei ole veel antud ravimit tarvitanud, võimalikest ohtudest klopidooreli manustamise korral.

Antud töös tuvastati WGS meetodi analüüsi abil ka teadaolematute sagedustega harva esinevaid allelele. *CYP2D6* geenil tuvastati kolm alleeli (*22, *83, *108), millele pole parimate meile teadaolevate andmete alusel üheski populatsioonis ja etnilises grupis täpseid sagedusi leitud ning samuti puudub informatsioon ka nende ennustatud fenotüübi kohta. Star-allelele *22 ja *83 on seni tuvastatud ainult aafriklastel/afroameeriklastel (Dodgen *et al.*, 2015; Gaedigk *et al.*, 2012). Alleelile *22 on määratud Dodgen *et al.*, 2015 põhjal PharmGKB sagedustetabelis esinemissageduseks 0,5% (www.pharmgkb.org/page/cyp2d6RefMaterials *CYP2D6* geeni spetsiifilised tabelid, külastatud 26. mai 2018). Star-alleleli *108 on varasemalt tuvastatud ainult ühes teadustöös Ameerikas ja ainult ühel indiviidil uuritud valimist, kelle etniline päritolu oli määramata (Qiao *et al.*, 2016). Kuigi antud allelele esineb <1% Eesti geenidoonoritest, võivad need alleelid olla olulise tähtsusega indiviidide metabolismi tüübi määramisel. Haruldased alleelid mõjutavad ligi 30-40% ulatuses varieeruvust ravimivastuses (Kozyra, *et al.*, 2017), mistõttu oleks oluline kasutada just põhjalikke sekveneerimis meetodeid, et määrata igale indiviidile võimalikult täpne fenotüüp, võttes arvesse ka harva esinevaid ja populatsioonispetsiifilisi variatsioone. Kuid teadaolevalt ja käesolevas töös ka kirjeldatud *CYP2D6* geenil on väga keerukas lookus ning umbes 95% ulatuses on selle varieeruvuse tuvastamine raskendatud, sest *CYP2D6*-ga paiknevad lähestikku ka kõrgelt homoloogsed geenid või pseudogeened, *CYP2D7* ja *CYP2D8*. WGS meetod, mis kasutab alleelide määramiseks lühikeste lugemite sekveneerimist (*short-read sequencing*), võib ebatäpselt kaardistada kõrge järjestuse homoloogiaga pseudogeene ning tulemuste analüüsimisel kattub mitu erinevat star-alleleli üksteisega (Fujikura *et al.*, 2015). Seetõttu on star-alleleli määramine raskendatud ning analüüsi tulemusena võib määrata indiviidile ebakorrekse genotüübi (Gaedigk, 2013). Mitme star-alleleli kattumine võib olla ka üheks põhjuseks, miks TÜ EGV täisgenoomi andmetest 2258-l oli diplotüüp määramatu. Näiteks, *CYP2D6**2 määramisel tuleb arvestada, et *CYP2D6**4 ja *CYP2D6**14 alleelide SNP-id võivad olla geneetiliselt aheldunud

star-alleeli *2 SNP-i variantide külge. Seega tuleks star-alleelide määramisel arvestada nii paljude erinevate SNP-ide järjestustega kui võimalik, et täpne genotüüp saada (Zhou *et al.*, 2017).

Geenil *CYP4F2* on senistes uurimistöödes tuvastatud kaks referentsalleelist erinevat star-alleeli, millel on kindlaks määratud SNP-id, vastavalt *2 ja *3 (www.pharmgkb.org/page/cyp4f2RefMaterials, külastatud 27. mai 2018). Kõige rohkem on uuritud star-alleeli *3. See *CYP4F2* alleel osaleb koos geenidega *CYP2C9* ja *VKORC1* antikoagulant varfariini metaboliseerimises. *CYP4F2* on vitamiin K oksüdaas, mis katalüüsib redutseeritud vitamiin K hüdroksü-vitamiin K1-ks ning eemaldab üleliigse vitamiin K metabolismirajast (Johnson *et al.*, 2017). Inimesed, kes on antud star-alleeli suhtes homosügootid, peaksid saama tavapärasest suurema annuse ravimit (Caldwell *et al.*, 2008; Borgiani *et al.*, 2009). CPIC suunistes on soovitatud *3 star-alleeliga patsientidel kõigis populatsioonides, v.a Aafrika, tõsta varfariini annust 5-10% võrra (Johnson *et al.*, 2017). Käesolevas töös leiti, et Eesti populatsioonis on *3 star-alleeli sagedus üsna madal. Enamikul geenidoonoritel määrati diplotüübiks *1/*3. Kuigi PharmGKB andmebaasis sellele diplotüübile fenotüübilist vastet ei ole antud, siis võib ennustada, et tegemist on tavapärasest aeglasemate metaboliseerijatega, sest neil on üks funktsionaalne alleel ja teine vähenenud aktiivsusega. Nendel 0,3% geenidoonoritest, kellel määrati diplotüüp *3/*3, oleks soovitatav aga edasi uurida, kas nad on või võiksid olla potentsiaalsed varfariini ravimi kasutajad. Antud informatsioon nende indiviidide diplotüübi kohta oleks tulevikus kasulik arstidele, kes siis saaksid vajadusel nendel patsientidel suurendada varfariini ravimiannust, et muuta ravi efektiivsemaks. Samuti leiti selles töös, et 63% geenidoonoritest on referentsalleelid ning normaalse metabolismiga. Paraku ei ole PharmGKB andmebaasis enamikele diplotüüpidele sagedusi euroopiidses populatsioonis veel määratud ning tulemuste kokkulangevust ei saa seetõttu omavahel võrrelda. Lisaks määrati TÜ EGV täisgenoomi andmetes 637 indiviidile ainult ühe alleeli star-alleel, kuid 46-le ei suudetud kumbagi alleeli määrata. Arvatakse, et *CYP4F2* geenil on tegelikult rohkem erinevaid SNP-e, kuid neid on veel vähe uuritud. Alleelide määramisel kasutatakse tavaliselt defineeritud ja teadaolevaid variatsioonide järjestusi. Sekveneerimine võimaldab identifitseerida ka uusi, seni tundmatuid alleeli variante, mis raskendavad genotüübi määramist. Kuigi nende variatsioonide mõju fenotüübile on raske hinnata, võivad need mõjutada näiteks *CYP4F2* rolli lipiidide (nt omega-3 rasvhapped) metabolismis (Kim *et al.*, 2018). Seetõttu võib arvata, et käesolevas töös TÜ EGV valimis leitud uus varieeruvuse kombinatsioon, millele ei vasta seni PharmGKB poolt kirjeldatud ja tuvastatud alleelid, võib olla uus alleel, mida seni pole lihtsalt veel identifitseeritud ning

fenotüüpi hinnatud. See osutab PharmGKB alleelide defineerimise tabelite nõrkustele, kuna kõik erinevates populatsioonides olemasolevad haplotüübid ei ole seal kaetud (ega saagi olla).

Teades geenidonorite *CYP* geenide kodeeritud ensüümide metabolismi fenotüüpe, on võimalik indiviididele jagada vajalikke ravimi soovitusi. Õige ravimiannuse määramine on kriitilise tähtsusega. Eriti oluline on see just NTI ravimite välja kirjutamisel. Põhiliselt puudutab see indiviide, kes on kas aeglased või ülikiired metaboliseerijad. Nii tagatakse efektiivne ravimi toime ja doos ning hoitakse potentsiaalselt ära ohtlikud kõrvaltoimed (Ahmed *et al.*, 2016). *CYP2C9* seos kahe ravimi, fenütoiin ja varfariin, metabolismis on kindlalt teaduslikult tõestatud. *CYP2C9* star-alleele *2 ja *3 loetakse vähenenud funktsiooniga alleelideks, mille olemasolu genoomis põhjustab ensüümi aktiivsuse langust (Caudle *et al.*, 2014; Johnson *et al.*, 2017). Indiviidid, kellele on määratud nende genotüübi põhjal keskmine (IM) või aeglane (PM) metabolismitüüp ning kellel on diagnoositud epilepsia, peaksid fenütoiini tavapärast ravimiannust vähendama, et vältida ADR-e, vastavalt 25% või 50% (Caudle *et al.*, 2014). TÜ EGV andmete põhjal on 635 geenidonoril *CYP2C9* fenotüübiks IM ning 60-l PM. Varfariini puhul seostatakse lisaks star-alleelidele *2 ja *3, varfariini metabolismi aktiivsuse vähenemist veel alleelidega *5, *6, *8 ja *11 (Johnson *et al.*, 2017), millest alleeli *11 tuvastati antud töös 0,1% indiviidil. Kirjeldatud informatsioon tavapärasest erinevate metabolismi fenotüüpide kohta oleks oluline teadmine TÜ EGV geenidonoritele, et varfariini ja fenotüüini manustamise korral kohandada raviplaan, vältimaks liiga efektiivset toimet või kõrvaltoimeid, milleks varfariini korral on risk verejooksudeks.

Geen *CYP3A5* reguleerib immuunsupressandi takroliimuse taset veres. TÜ EGV andmetes oli domineerivaimaks fenotüübiks PM, mis on ka eelistatud takroliimuse metabolismis. *CYP3A5**3 põhjustab alternatiivset splaissingut, mille tõttu kaotab sünteesitud valk oma aktiivsuse. Defektset *CYP3A5**3 esineb kõige sagedamini Euroopa populatsioonis, 94,3% (Zhou *et al.*, 2017). Ravimi metabolismis on see fenotüüp eelistatud. Aeglasem metabolism aitab hoida veres püsivamat immuunsupressandi takroliimuse taset, mis aitab ära hoida organi siirdamisel äratõukereaktsiooni (Birdwell *et al.*, 2015). Seega need indiviidid ei vaja tavapärase annuse suurendamist või vähendamist. Geenidonoritest 327-l määrati fenotüübiks aga IM. Indiviidid IM fenotüübiga, kellele on arstid kirjutanud takroliimust, on soovitatav võtta tavapärasest 1,5-2 korda suuremat ravimiannust (Birdwell *et al.*, 2015).

TÜ EGV geenidonorite seas on *CYP2D6* fenotüüpidest kõige sagedasem tavapärasest aeglasem metabolism (PM), mida kõige sagedamini põhjustab nii Eesti geenidonorite seas kui ka

europiidises populatsioonis (Sistonen *et al.*, 2007) mittefunktsionaalne star-alleel *4. *CYP2D6**4 variatsiooni tulemusena toimub vigane splaissing ning tulemuseks on inaktiveeritud *CYP2D6* geeni produkt (Zhou *et al.*, 2017). See fenotüüp määrab ära mitme ravimi saatuse organismis. Näiteks, kodeiini, mis on opioid, mille *CYP2D6* ensüümid muudavad organismis morfiiniks. Indiviididel, aga, kellel on fenotüüp PM või UM, on soovitatav kasutada alternatiivseid valuvaigisteid, et vähendada võimalikke ebameeldivaid kõrvalmõjusid ning ravimi toksilisust (Crews *et al.*, 2014). Samuti peaksid PM ja UM fenotüübiga patsiendid valima raviks alternatiivseid ravimeid tritsüklilistele antidepressantidele (nt amitriptüliin) või reguleerima ravimiannust (Hicks *et al.*, 2017).

CYP2D6 genotüpiseerimist, sekveneerimist ja fenotüpiseerimist mõjutavad ka geeni CNV-d. Selleks et võimalikult täpselt määrata *CYP2D6* koopiaarve, tuleks kasutada selliseid DNA analüüsi sonde, mis on suunatud erinevatele piirkondadele *CYP2D6* geenis (nt ekson 9, intron 2, intron 6). Nii saab lisaks koopiaarvule kindlaks teha ka hübriidgeenide ja nende tandemite täpset paiknemist ja kirjeldust, mis võimaldaks täpsemalt määrata indiviidi metabolismi fenotüüpi (Gaedigk, 2013). Ekson 9 analüüs suudab tuvastada kõiki allelele, välja arvatud mittefunktsionaalset *CYP2D6**36 star-alleleli. Intron 2 ja 6 suudavad tuvastada ka star-alleleli *36 (Ramamoorthy *et al.*, 2010). Ühes kromosoomis võib olla ühe alleleli asemel ka kaks erinevat alleleli, mis moodustavad tandemeid, näiteks *36+*10, mida esineb sagedamini Aasia populatsioonis. Kui ei suudeta tuvastada tandemseid allelele, siis võidakse seda tõlgendada kui geeni duplikatsiooni ning indiviidile määratakse vale fenotüüp (Yang *et al.*, 2017). Seetõttu on oluline *CYP2D6* uurimisel ning vastava fenotüübi ennustamisel määrata nii järjestuse variatsioone kui ka koopiaarvu erinevusi. Ainult genotüpiseerimise tulemustest ei piisa, et arstid haiglates saaksid kirjutada patsientidele suunatud, individuaalsete soovitustega ravi. Täpne ja korrektne genotüüp-fenotüüp määramine on kõrgelt hinnatud abivahend haiglates, kus see aitab juba praegu, kuid kindlasti tulevikus aina enam, arstidel ravimääramisel patsientidele kirjutada individuaalsete suunatud soovitustega ravi.

Käesolevas bakalaureusetöös analüüsiti ja võrreldi kahte arvutuslikku meetodit, kus määrati *CYP2D6* geenis koopiaarvud kolme regiooni alusel, vastavalt 22:42523514-42526507, 22:42525259-42531614, 22:42537191-42540184. Valideerimiseks kasutati *TaqMan* analüüsi ekson 9 ja intron 2 piirkondi. Valideerimise tulemusena järeldus, et võrreldes *k*-meer meetodiga kattuvad *TaqMan* analüüsi tulemused suuremal määral *Genome STRiP* meetodiga. Kuigi *Genome STRiP* ja *k*-meer meetodite tulemuste valideerimiseks kasutatava *TaqMan* analüüsi valimisse olid

valitud ainult 17 geenidoonori proovid kõigist 230-st arvutuslike meetodite vahel tuvastatud koopiarvude erinevustest, siis nende 17 proovi *TaqMan* koopiarvud kattusid täielikult *Genome STRiP*’iga. Esimese arvutusliku meetodi ja *TaqMan* analüüside tulemused on tugevas lineaarses korrelatsioonis. Eelnevate tulemuste põhjal võib järeldada, et antud analüüsis on *Genome STRiP* määratud koopiarvud usaldusväärsemad kui *k*-meeride meetodi. *TaqMan* meetodeid loetakse üldiselt täpsemaks kui arvutuslike meetodeid, sest nende suurimaks eeliseks on võime identifitseerida ka *CYP2D6* pseudogeenide olemasolu. Teisalt võivad nii analüüsimeetodi vead kui ka andmete varieeruvus viia valepositiivsete tulemusteni. Arvutuslikud meetodid on vastupidiselt aga kiiremad ja lihtsamad. Vaatamata sellele andsid arvutuslikud meetodid suuremaid koopiarve (>4) sagedamini kui *TaqMan* analüüs. Hetkel ei ole veel täpselt selge, millest see erinevus võib tulla. Ühe variandina võib pakkuda *CYP2D6* keerulist lookust ning pseudogeenide olemasolu, mida on raske detekteerida. Näiteks, *k*-meer meetodi valikusse sattusid need *k*-meerid, mis olid referentsjärjestuses vaid ühes korduses (unikaalsed *k*-meerid) ehk nende varieeruvad positsioonid olid *CYP2D6* põhised. Samuti võiks kõrgemat koopiarvu seletada arvutuslike ebatäpsustega. *K*-meer meetodis kasutati koopiarvude arvutamise algoritmides referentsgenoomi pikkust. Tegelikult varieerub iga indiviidi genoomi pikkus ning see võib kohati mõjutada valepositiivsete tulemuste saamise tõenäosust (Auton *et al.*, 2015). Seega ei suuda arvutuslikud meetodid veel täielikult katselist määramist asendada. Koopiarvude määramisel on kindlasti oluline valida sobiv meetod, mis arvestades populatsiooni iseärasusi ning uuringu eesmärki, annab kõige usaldusväärsema tulemuse. Populatsioonides, kus esineb rohkem *CYP2D6**36 alleeli, nt Aasias, on soovitatav kasutada *TaqMan* ekson 9 meetodile alternatiivseid analüüse, sest see ei tuvasta vastavat star-alleeli. Soovitatud on ka kombineerida erinevaid CNV-sid detekteerivaid *TaqMan* analüüse populatsioonides, kus võib esineda star-alleeli *36. Näiteks, koos ekson 9 analüüsiga kasutada ka kas intron 2 või intron 6 regiooni analüüsi, mis mõlemad tuvastavad *36 alleeli (Ramamoorthy *et al.*, 2010).

Eesti populatsioonis on oluline uurida *CYP* geeniperekonna allelele, sest antud töös selgus, et tavapärasest metabolismi tüübist erinev fenotüüp, mis vajaks CPIC suuniseid ravimi annuse reguleerimiseks, esineb umbes 47% geenidoonoritest vähemalt ühes *CYP* geenis viiest analüüsitud geenist. Üldine pilt geenidoonorite alleelide esinemissagedustest ja koopiarvudest annab olulist informatsiooni edaspidiste SNP-ide ja CNV-de analüüside ette valmistamiseks ja läbi viimiseks.

KOKKUVÕTE

Bakalaureusetöö uuris, milline on ravimite metabolismi reguleerivate *CYP* geenide esinemissagedus ning ennustatav fenotüüp TÜ EGV doonorite hulgas ning kas ja millisel määral erineb see europiidses populatsioonis leitud sagedustest. Samuti oli antud töö eesmärgiks hinnata *CYP2D6* geeni koopiaarvude varieeruvust Eesti geenidoonorite seas, kasutades võrdlemiseks kolme erinevat meetodikat (*Genome STRiP*, *k-meerid*, *TaqMan*) koopiaarvude leidmiseks.

Töös leitud sagedused langesid suures osas kokku europiidsele populatsioonile seni leitud alleeli sagedustega. Kuigi 35,9% geenidoonoritest kannab geenis *CYP2C19* diplotüüpi $*1/*1$, siis umbes samal suurel hulgal (31,2%) doonoritest määrati diplotüübiks $*1/*17$, mis märgib tavapärasest kiiremat metabolismitüüpi. Europiidses populatsioonis on määratud diplotüübile $*1/*17$ kaks korda väiksem sagedus. Nii selle diplotüübi suurele sagedusele kui ka olulisus ravimi metabolismis näitab, et on oluline määrata *CYP2C19* alleelide esinemissagedusi. Lisaks leiti töös, et *CYP2D6* ja *CYP4F2* diplotüüpide määramine on raskendatud ning leiti ka sagedusi haruldastele alleelidele. *CYP4F2* korral tuvastati 28,2% sagedusega esinev potentsiaalselt uus alleel. Paraku on *CYP4F2* alleelide esinemist europiidses populatsioonis veel üsna vähe uuritud, nii et geenidoonoritel määratud diplotüüpide sageduste võrdlemine PharmGKB andmebaasiga ei olnud võimalik

CYP2D6 korrektse fenotüübi määramiseks on lisaks geeni järjestuste kindlaks tegemisele oluline määrata ka geeni koopiaarve. Käesolevas töös võrreldi kahte arvutuslikku koopiaarvude määramise meetodikat meetodikat ja valideerimiseks kasutati relatiivset kvantiseerimist *TaqMan Copy Number Assay* ga. Kõigi meetodite tulemused olid tugevas lineaarses korrelatsioonis. Ülekaalukalt leiti TÜ EGV doonoritel *CYP2D6* geeni koopiaarvuks 2, mis märgib funktsionaalset geeni koopiat.

Käesolevas töös teostatud eksperimentaalse osa tulemusena järeldub, et TÜ EGV täisgenoomi andmete alusel esineb umbes 47% analüüsitud geenidoonoritest tavapärasest erineva metabolismiga fenotüüp vähemalt ühes viiest analüüsitud *CYP* geenist. Tulevikus võiks indiviidide genotüüp-fenotüüp paaride info olla kättesaadav ka arstidele, arendades nii personaalset meditsiini. Sel moel oleks võimalik tuvastada patsiente, kellel on kõrge riskiga genotüübid (PM või UM) ravimite lagundamises ning määrata neile personaalsete soovitusetega ravi, mis võiks tõsta ravi efektiivsust ning vähendada ebameeldivusi patsiendile.

Identification of genetic variation in drug metabolizing cytochrome P450 genes in Estonian population

Karin Kõrgesaar

SUMMARY

Different patients can respond differently to the same drug and dose. Sometimes, the effective drug dose for a particular patient may prove lethal or result in therapeutic failure in others (too low drug concentrations at normal doses), leading to serious adverse drug reactions (ADR) or no effects at all. Interindividual variation of drug response is caused by a combination of genetic and environmental factors (nutrition, lifestyle, chemical exposure, drug-drug interactions, etc.).

Cytochrome P450 (CYP) is a large enzyme superfamily that regulates phase I drug metabolism. Genetic polymorphism, such as single nucleotide polymorphism (SNP), copy number variations (CNV), gene deletions and duplications, in genes encoding CYP enzymes, affects significantly phenotypic drug responses. Based on the variability in CYP genes, individuals can be classified into 4 major phenotypes: ultrarapid metabolizer (UM) with more than two functional alleles, rapid metabolizer (RM), who carry at least one increased function allele, extensive metabolizer (EM) carrying two normally-functioning alleles, intermediate metabolizer (IM) usually carrying 1 functional and 1 defective allele and poor metabolizer (PM) with two non-functional alleles. Since the advent of high throughput genotyping and next-generation sequencing (NGS) technologies, numerous studies all over the world have identified many important allelic variants associated with drug response in *CYP* genes. As a result, CPIC and PharmGKB have been founded to give out guidelines that enable the translation of genetic laboratory test results into actionable prescribing decisions for specific drugs. This knowledge helps to make drug therapy more effective and reduce adverse reactions.

Strikingly, the distribution of *CYP* alleles differs considerably between populations. Therefore, the aim for this study was to derive the frequencies of alleles and predicted phenotypes of five different *CYP* genes (*CYP2C9*, *CYP2C19*, *CYP2D6*, *CYP3A5*, *CYP4F2*) in Estonian population, and compare the results with the frequencies in PharmGKB found across Caucasians. Moreover, *CYP2D6*, which contributes to the metabolism of up to 25% of clinically used drugs, has a very complex loci, thereby posing problems in successful sequencing by short-read sequencing projects.

For that reason, it is important to determine the gene copy number for accurate and informed genotype interpretation. So, in this research the results of four different assays for detecting CNV in *CYP2D6* were also analysed, and correlation was identified between three of them.

This study is based on the data of University of Tartu Estonian Genome Centre. Frequencies of *CYP* genes were determined in 2420 gene donors. To assess the *CYP2D6* copy number, four assays were used: Genome STRiP pipeline (n=2269), *k*-mer-based method (n=1873), which target three different regions of the gene (22:42523514-42526507, 22:42525259-42531614, 22:42537191-42540184) and two TaqMan Copy Number Assays (n=230), that target intron 2 and exon 9.

The frequencies of alleles of *CYP* genes do not vary considerably between Estonian population and Caucasians. Nevertheless, some rare variations were found. Genetic variability of *CYP2C9* is dominated by the diplotype **1/*1*, that is associated with extensive metabolism. *CYP2C19* shows the highest diplotype frequencies for **1/*1* (35.9%) and **1/*17* (31.2%), respectively. Notably, the frequency of **1/*17* found across Estonian gene donors, is almost two times higher than stated in PharmGKB for Caucasians. This diplotype causes increased enzyme activity and is associated with rapid metabolism. As expected, 85.8% of gene donors are carrying two non-functional alleles **3/*3* in *CYP3A5*. These individuals are *CYP3A5* nonexpressers (poor metabolizers). When it comes to the drug metabolism, PM phenotype is favoured and helps to keep the target drug concentration in the blood. Even though genes *CYP2D6* and *CYP4F2* presented the most abundant diplotypes found in Caucasians (*CYP2D6***4/*4* and *CYP4F2***1/*1*), there were individuals, whose alleles were undetermined. These results suggest that undetermined alleles could be rare variants, which have not been studied before or not included in this testing platform. However, may have important effects on drug response.

The estimated *CYP2D6* gene copy numbers from the four assays were concordant in most of the samples. The most frequent copy number found was two, which is functional gene copy. In addition, a positive linear correlation between Genome STRiP pipeline and TaqMan assays was identified.

The results from this study indicate the importance of detecting individual's genotype and translate it into phenotype. Based on the findings in this paper, more than 47% of gene donors have phenotypes different from normal metabolizer in at least one of the *CYP* genes that influence drug treatment in many different drugs (e.g. antidepressants, opioids, anticoagulant etc.). Therefore, this knowledge helps patients to receive efficient drug therapy and avoid adverse reactions.

TÄNUAVALDUSED

Soovin tänada eelkõige oma juhendajaid Kristi Krebsi ja Lili Milanit võimaluse eest teha oma bakalaureusetöö TÜ EGV farmakogeneetika töögrupis. Samuti tahaksin tänada oma juhendajaid toetuse ja nõuannete eest töö koostamise protsessis.

KASUTATUD KIRJANDUS

Ahmed, S., Zhou, Z., Zhou, J. & Chen, S. Q. (2016). Pharmacogenomics of Drug Metabolizing Enzymes and Transporters: Relevance to Precision Medicine. *Genomics, Proteomics & Bioinformatics* 14 (5): 298–313.

Alving, A. S., Carson, P. E., Flanagan, C. L. & Ickes, C. E. (1956). Enzymatic Deficiency in Primaquine-Sensitive Erythrocytes. *Science (New York, N.Y.)*, 124 (3220): 484-5.

Arbitrio, M., Di Martino, M. T., Scionti, F., Agapito, G., Guzzi, P. H., Cannataro, M., Tassone, P. & Tagliaferri, P. (2016). Drug Metabolism Enzymes and Transporters: A Pharmacogenomic Platform for Precision Medicine. *Oncotarget*, 7 (33): 54028–50.

Auton, A., Brooks, L. D., Durbin, R. M., ... Abecasis, G. R. (2015). A Global Reference for Human Genetic Variation. *Nature*, 526 (7571): 68-74.

Birdwell, K. A., Decker, B., Barbarino, J. M., ... MacPhee, I. A. (2015). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guidelines for *CYP3A5* Genotype and Tacrolimus Dosing. *Clinical Pharmacology and Therapeutics*, 98 (1): 19–24.

Blix, H. S., Viktil, K. K., Moger, T. A. & Reikvam, A. (2010). Drugs with Narrow Therapeutic Index as Indicators in the Risk Management of Hospitalised Patients. *Pharmacy Practice* 8 (1): 50–55.

Borgiani, P., Ciccacci, C., Forte, V., Sirianni, E., Novelli, L., Bramanti, P. & Novelli, G. (2009). *CYP4F2* Genetic Variant (Rs2108622) Significantly Contributes to Warfarin Dosing Variability in the Italian Population. *Pharmacogenomics*, 10 (2): 261–66.

Caldwell, M. D., Awad, T., Johnson, J. A., ... Burmester, J. K. (2008). *CYP4F2* Genetic Variant Alters Required Warfarin Dose. *Blood*, 111 (8): 4106–12.

Caudle, K. E., Rettie, A. E., Whirl-Carrillo, M., Smith, L. H., Mintzer, S., Lee, M. T., Klein, T. E. & Callaghan, J. T. (2014). Clinical Pharmacogenetics Implementation Consortium Guidelines for *CYP2C9* and HLA-B Genotypes and Phenytoin Dosing. *Clinical Pharmacology and Therapeutics*, 96 (5): 542-8.

Crews, K. R., Gaedigk, A., Dunnenberger, H. M., ... Skaar, T. C. (2014). Clinical Pharmacogenetics Implementation Consortium Guidelines for Cytochrome P450 2D6 Genotype and Codeine Therapy: 2014 Update. *Clinical Pharmacology and Therapeutics*, 95 (4): 376–82.

Daly, A. K. (2010). Genome-Wide Association Studies in Pharmacogenomics. *Nature Reviews Genetics*, 11 (4): 241–46.

Dodgen, T. M., Eloff, A., Mataboge, C., Roos, L. J., Van Staden, W. C. & Pepper, M. S. (2015). Risperidone-Associated Adverse Drug Reactions and *CYP2D6* Polymorphisms in a South African Cohort. *Applied & Translational Genomics*, 14 (5): 40-6.

Doogue, M. P. & Polasek, T. M. (2013). The ABCD of Clinical Pharmacokinetics. *Therapeutic Advances in Drug Safety*, 4 (1): 5-7.

- Fujikura, K., Ingelman-Sundberg, M. & Lauschke, V. M. (2015). Genetic Variation in the Human Cytochrome P450 Supergene Family. *Pharmacogenetics and Genomics*, 25 (12): 584–94.
- Gaedigk, A. (2013). Complexities of *CYP2D6* Gene Analysis and Interpretation. *International Review of Psychiatry*, 25 (5): 534–53.
- Gaedigk, A., Sangkuhl, K., Whirl-Carrillo, M., Klein, T. & Leeder, J. S. (2017). Prediction of *CYP2D6* Phenotype from Genotype across World Populations. *Genetics in Medicine*, 19 (1): 69–76.
- Gaedigk, A., Twist, G. P. & Leeder, J. S. (2012). *CYP2D6*, *SULT1A1* and *UGT2B17* Copy Number Variation: quantitative detection by multiplex PCR. *Pharmacogenomics*, 13 (1): 91–111.
- Gonzalez, F., Skodak, R. C., Kimura, S., Umeno, M., Zanger, U. M., Nebert, D. W., Gelboin, H. V., Hardwick, J. P. & Meyer, U. A. (1988). Characterization of the Common Genetic Defect in Humans Deficient in Debrisoquine Metabolism. *Nature*, 331 (6155): 442–46.
- Gopisankar, M. G. (2017). *CYP2D6* Pharmacogenomics. *Egyptian Journal of Medical Human Genetics*, 18 (4): 309–13.
- Handsaker, R. E., Van Doren, V., Berman, J. R., Genovese, G., Kashin, S., Boettger, L. M. & McCarroll, S. A. (2015). Large Multiallelic Copy Number Variations in Humans. *Nature Genetics*, 47 (3): 296–303.
- Hicks, J. K., Sangkuhl, K., Swen, J. J., ... Stingl, J. C. (2017). Clinical Pharmacogenetics Implementation Consortium Guideline (CPIC) for *CYP2D6* and *CYP2C19* Genotypes and Dosing of Tricyclic Antidepressants: 2016 Update. *Clinical Pharmacology & Therapeutics*, 102 (1): 37–44.
- Ingelman-Sundberg, M., Sim, S. C., Gomez, A. & Rodriguez-Antona, C. (2007). Influence of Cytochrome P450 Polymorphism on Drug Therapies: pharmacogenetic, pharmacoepigenetic and clinical aspects. *Pharmacology & Therapeutics*, 116 (3): 496–526.
- Jiang, W., Makhoul, F., Schuirmann, D. J., Zhang, X., Zheng, N., Conner, D., Yu, L. X. & Lionberger, R. (2015). A Bioequivalence Approach for Generic Narrow Therapeutic Index Drugs: Evaluation of the Reference-Scaled Approach and Variability Comparison Criterion. *The AAPS Journal*, 17 (4): 891–901.
- Johnson, J. A., Caudle, K. E., Gong, L., ... Wadelius, M. (2017). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for Pharmacogenetics-Guided Warfarin Dosing: 2017 Update. *Clinical Pharmacology and Therapeutics*, 102 (3): 397–404.
- Kalman, L. V., Agúndez, J., Appell, M. L., ... Zanger, U. M. (2016). Pharmacogenetic Allele Nomenclature: International Workgroup Recommendations for Test Result Reporting. *Clinical Pharmacology and Therapeutics*, 99 (2): 172–85.
- Kalow, W. (2006). Pharmacogenetics and Pharmacogenomics: Origin, Status, and the Hope for Personalized Medicine. *The Pharmacogenomics Journal*, 6 (3): 162–65.

- Kaplinski, L., Lepamets, M. & Remm, M. (2015). GenomeTester4: A Toolkit for Performing Basic Set Operations - Union, Intersection and Complement on k-Mer Lists. *GigaScience*, 4: 58.
- Kim, W. Y., Lee, S. J., Min, J., Oh, K. S., Kim, D. H., Kim, H. S. & Shin, J. G. (2018). Identification of Novel *CYP4F2* Genetic Variants Exhibiting Decreased Catalytic Activity in the Conversion of Arachidonic Acid to 20-Hydroxyeicosatetraenoic Acid (20-HETE). *Prostaglandins, Leukotrienes, and Essential Fatty Acids*, 131 (4): 6–13.
- Kimura, S., Umeno, M., Skoda, R. C., Meyer, U. A. & Gonzalez, F. J. (1989). The Human Debrisoquine 4-Hydroxylase (CYP2D) Locus: Sequence and Identification of the Polymorphic *CYP2D6* Gene, a Related Gene, and a Pseudogene. *American Journal of Human Genetics*, 45 (6): 889–904.
- Kozyra, M., Ingelman-Sundberg, M. & Lauschke, V. M. (2017). Rare Genetic Variants in Cellular Transporters, Metabolic Enzymes and Nuclear Receptors Can Be Important Determinants of Interindividual Differences in Drug Response. *Genetics in Medicine*, 19 (1): 20–29.
- Maggo, S. D., S., Savage, R. L. & Kennedy, M. A. (2016). Impact of New Genomic Technologies on Understanding Adverse Drug Reactions. *Clinical Pharmacokinetics*, 55 (4): 419–36.
- Mahgoub, A., Idle, J. R., Dring, L. G., Lancaster, R. & Smith, R. L. (1977). Polymorphic Hydroxylation of Debrisoquine in Man. *The Lancet*, 310 (8038): 584–86.
- Marez, D., Legrand, M., Sabbagh, N., Lo Guidice, J. M., Spire, C., Lafitte, J. J., Meyer, U. A. & Broly, F. (1997). Polymorphism of the Cytochrome P450 *CYP2D6* Gene in a European Population: Characterization of 48 Mutations and 53 Alleles, Their Frequencies and Evolution. *Pharmacogenetics*, 7 (3): 193–202.
- McDonagh, E. M., Whirl-Carrillo, M., Garten, Y., Altman, R. B. & Klein, T. E. (2011). From Pharmacogenomic Knowledge Acquisition to Clinical Applications: The PharmGKB as a Clinical Pharmacogenomic Biomarker Resource. *Biomarkers in Medicine*, 5 (6): 795–806.
- Meyer, U. A. (2004). Pharmacogenetics - Five Decades of Therapeutic Lessons from Genetic Diversity. *Nature Reviews Genetics*, 5 (9): 669–76.
- Motulsky, A. G. (1957). Drug Reactions Enzymes, and Biochemical Genetics. *Journal of the American Medical Association*, 165 (7): 835–37.
- Ozdemir, V., Kashuba, A. D. M., Basile, V. S. & Kennedy, J. L. 2002. Pharmacogenetics of Psychotropic Drug Metabolism, p. 157-80. *In* B. Lerer (ed.), *Pharmacogenetics of Psychotropic Drugs*, Cambridge University Press, Cambridge.
- Pedersen, R. S., Brasch-Andersen, C., Sim, S. C., ... Ingelman-Sundberg, M. (2010). Linkage Disequilibrium between the *CYP2C19*17* Allele and Wildtype *CYP2C8* and *CYP2C9* Alleles: Identification of CYP2C Haplotypes in Healthy Nordic Populations. *European Journal of Clinical Pharmacology*, 66 (12): 1199–1205.
- Pirmohamed, M. (2014). Personalized Pharmacogenomics: Predicting Efficacy and Adverse Drug Reactions. *Annual Review of Genomics and Human Genetics*, 15 (1): 349–70.

- Qiao, W., Yang, Y., Sebra, R., Mendiratta, G., Gaedigk, A., Desnick, R. J. & Scott, S. A. (2016). Long-Read Single Molecule Real-Time Full Gene Sequencing of Cytochrome P450-2D6. *Human Mutation*, 37 (3): 315-23.
- Ramamoorthy, A., Flockhart, D. A., Hosono, N., Kubo, M., Nakamura, Y. & Skaar, T. C. (2010). Differential Quantification of *CYP2D6* Gene Copy Number by Four Different Quantitative Real-Time PCR Assays. *Pharmacogenetics and Genomics*, 20 (7): 1.
- Rowland, A., Miners, J. O. & Mackenzie, P. I. (2013). The UDP-glucuronosyltransferases: Their Role in Drug Metabolism and Detoxification. *The International Journal of Biochemistry & Cell Biology*, 45 (6): 1121-32.
- Sachse, C., Brockmöller, J., Bauer, S. & Roots, I. (1997). Cytochrome P450 2D6 Variants in a Caucasian Population: Allele Frequencies and Phenotypic Consequences. *American Journal of Human Genetics*, 60 (2): 284–95.
- Scott, S. A., Sangkuhl, K., Stein, C. M., Hulot, J. S., Mega, J. L., Roden, D. M., Klein, T. E., Sabatine, M. S., Johnson, J. A. & Shuldiner, A. R. (2013). Clinical Pharmacogenetics Implementation Consortium Guidelines for *CYP2C19* Genotype and Clopidogrel Therapy: 2013 Update.
- Shirasaka, Y., Chaudhry, A. S., McDonald, M., ... Thummel, K. E. (2016). Interindividual Variability of *CYP2C19*-Catalyzed Drug Metabolism Due to Differences in Gene Diplotypes and Cytochrome P450 Oxidoreductase Content. *Pharmacogenomics Journal*, 16 (4): 375–87.
- Sibbing, D., Koch, W., Gebhard, D., Schuster, T., Braun, S., Stegherr, J., Morath, T., Schomig, A., Von Beckerath, N. & Kastrati, A. (2010). Cytochrome 2C19*17 Allelic Variant, Platelet Aggregation, Bleeding Events, and Stent Thrombosis in Clopidogrel-Treated Patients With Coronary Stent Placement. *Circulation*, 121 (4): 512–18.
- Sim, S. C., Daly, A. K. & Gaedigk, A. (2012). *CYP2D6* Update: Revised Nomenclature for *CYP2D7/2D6* Hybrid Genes. *Pharmacogenetics and Genomics*, 22: 692–94.
- Sistonen, J., Sajantila, A., Lao, O., Corander, J., Barbuji, G. & Fuselli, S. (2007). *CYP2D6* Worldwide Genetic Variation Shows High Frequency of Altered Activity Variants and No Continental Structure. *Pharmacogenetics and Genomics*, 17 (2): 93–101.
- Swen, J J, M Nijenhuis, A de Boer, L Grandia, A H Maitland-van der Zee, H Mulder, G A P J M Rongen, et al. 2011. “Pharmacogenetics: From Bench to Byte— An Update of Guidelines.” *Clinical Pharmacology & Therapeutics* 89 (5): 662–73. <https://doi.org/10.1038/clpt.2011.34>.
- Sychev, D. A., Denisenko, N. P., Sizova, Z. M., Grachev, A. V. & Velikolug, K. A. (2015). The Frequency of *CYP2C19* Genetic Polymorphisms in Russian Patients with Peptic Ulcers Treated with Proton Pump Inhibitors. *Pharmacogenomics and Personalized Medicine*, 8: 111–14.

Watson, J. D. & Crick, F. H. C. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature*, 171 (4356): 737–38.

Whirl-Carrillo, M., McDonogh, E. M., Herbet, J., Gong, L., Sangkuhl, K., Thotn, C., Altman, R. & Klein, E. (2012). Pharmacogenomics Knowledge for Personalized Medicine. *Clinical Pharmacology and Therapeutics*, 92 (4): 414–17.

Yang, Y., Botton, M. R., Scott, E. R. & Scott, S. A. (2017). Sequencing the *CYP2D6* Gene: From Variant Allele Discovery to Clinical Pharmacogenetic Testing. *Pharmacogenomics*, 18 (7): 673–85.

Yiannakopoulou, E. Ch. (2013). Pharmacogenomics of Phase II Metabolizing Enzymes and Drug Transporters: Clinical Implications. *The Pharmacogenomics Journal*, 13 (2): 105-109.

Zanger, U. M. & Schwab, M. (2013). Cytochrome P450 Enzymes in Drug Metabolism: Regulation of Gene Expression, Enzyme Activities, and Impact of Genetic Variation. *Pharmacology & Therapeutics*, 138 (1): 103–41.

Zawilska, J. B., Wojcieszak, J. & Olejniczak, A. B. (2013). Prodrugs: A Challenge for the Drug Development. *Pharmacological Reports: PR* 65 (1): 1–14.

Zhou, S. F., (2009). Polymorphism of Human Cytochrome P450 2D6 and Its Clinical Significance: Part I. *Clinical Pharmacokinetics*, 48 (12): 761-804.

Zhou, S. F., Di, Y. M., Chan, E., ... Duan, W. (2008). Clinical Pharmacogenetics and Potential Application in Personalized Medicine. *Current Drug Metabolism*, 9 (8): 738-84.

Zhou, Y., Ingelman-Sundberg, M. & Lauschke, V. M. (2017). Worldwide Distribution of Cytochrome P450 Alleles: A Meta-Analysis of Population-Scale Sequencing Projects. *Clinical Pharmacology and Therapeutics*, 102 (4): 688–700.

KASUTATUD VEEBIAADRESSID

1. <https://cpicpgx.org/>
2. <https://ghr.nlm.nih.gov/>
3. <https://genome.ucsc.edu/>
4. <https://www.pharmgkb.org/>
5. <https://www.thermofisher.com>

LISAD

Lisa 1. Kahe arvutusliku meetodi, *Genome STRiP* ja *K-meerid*, erinevused koopiaarvude määramisel ning *TaqMan* ekson 9 ja intron 2 analüüside kokkulangevus leitud koopiaarvudega samadele indiviididele.

Genome STRiP	K-meer	Ekson 9	Intron 2
2	3	2	2
2	1	2	2
2	1	2	2
2	1	2	2
3	2	3	3
2	3	2	2
3	1	3	4
2	1	2	2
3	2	3	3
3	4	3	4
2	1	2	2
2	1	2	2
3	2	3	3
3	4	3	4
2	1	2	2
2	3	2	2
2	1	2	2

Lisa 2. 96 geenidoonori andmed, kellel leitud koopiaarv klappis nii *Genome STRiP*, *TaqMan* ekson 9 kui ka *TaqMan* intron 2 meetodiga ning kellel suudeti määrata *CYP2D6* diplotüüp.

Koopiaarv	Diplotüüp
1	*4/*4
1	*4/*4
1	*4/*4
1	*4/*4
1	*4/*4
2	*4/*33
2	*9/*33
2	*4/*4
2	*3/*4
2	*4/*4
2	*4/*33
2	*3/*4
2	*3/*4
2	*4/*4
2	*4/*4
2	*4/*4

Koopiaarv	Diplotüüp
2	*4/*4
2	*4/*4
2	*4/*4
2	*4/*6
2	*4/*33
2	*4/*9
2	*4/*4
2	*4/*4
2	*4/*4
2	*4/*9
2	*4/*4
2	*4/*4
2	*3/*4
2	*3/*4
2	*3/*4
2	*4/*4

Lisa 3. Koopiaarvude erinevused TÛ EGV andmetest *Genome STRiP* (GS) meetodi ja *TaqMan* ekson 9 vahel.

GS	Exon 9	Intron 2	<i>CYP2D6</i>
2	1	2	*4/-
2	1	2	Määramatu
1	2	1	Määramatu
1	2	1	*3/*3
1	2	1	Määramatu
1	2	1	*4/*4
3	2	3	*33/-
3	2	4	Määramatu
2	3	2	Määramatu
3	2	2	*3/*4

Lisa 4. Koopiaarvude erinevused 23 indiviidil TÜ EGV andmetest *Genome STRiP* (GS) meetodi ja *TaqMan* intron 2 vahel.

GS	Intron 2	Ekson 9	<i>CYP2D6</i>
2	1	2	*4/-
2	1	2	*4/*4
2	1	2	*4/*4
2	3	2	*4/-
2	3	2	*4/-
2	3	2	*4/-
2	3	2	*4/*4
2	3	2	-/*4
2	3	2	*4/*4
2	3	2	*4/-
2	3	2	Määramatu
2	3	2	*4/-
2	3	2	*4/-
2	3	2	*4/-
3	4	2	Määramatu
3	2	3	Määramatu
3	2	3	-/*4
3	4	3	Määramatu
3	4	3	*4/-
3	4	3	-/*6
3	4	3	Määramatu
4	5	4	Määramatu
3	2	2	*3/*4

LIHTLITSENTS

Lihlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Karin Kõrgesaar (sünnikuupäev: 12.12.1993)

1. annan Tartu Ülikoolile tasuta loa (lihlitsentsi) enda loodud teose Ravimivastuses oluliste tsütokroom P450 geenide varieeruvuse määramine Eesti populatsioonis, mille juhendajad on Kristi Krebs ja Lili Milani,
 - 1.1.reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
 - 1.2.üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu alates 06.04.2018 kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et nimetatud õigused jäävad alles ka autorile.
3. kinnitan, et lihlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 28.05.2018